

**ALMA MATER STUDIORUM – UNIVERSITA' DI
BOLOGNA**

SCUOLA DI ECONOMIA E MANAGEMENT

Corso di Laurea Magistrale in Economia e Politica Economica

**STADI DI SVILUPPO DELLE IMPORTAZIONI:
UN'ANALISI EMPIRICA**

Econometria Avanzata

Presentata da:

nome e cognome: Alessandro Ferrari

matricola: 943615

Relatore:

Prof. Roberto Golinelli

APPELLO IV Marzo

ANNO ACCADEMICO 2020 / 2021

Indice

Introduzione	5
1 Un quadro teorico di riferimento	7
1.1 Fasi di Sviluppo secondo Rostow	7
1.1.1 Società Tradizionale	8
1.1.2 Precondizioni per il decollo industriale	8
1.1.3 Decollo Industriale	10
1.1.4 Maturità	10
1.1.5 Era del consumo di massa	11
1.2 Ciclo vitale di un Prodotto	12
1.2.1 Che cos'è il ciclo vitale di un'impresa o di un settore? .	12
1.2.2 Le quattro fasi del ciclo vitale del prodotto	13
1.2.3 Concetto di Business Risk	15
1.3 Cicli di apertura di un paese alle importazioni di un prodotto	17
1.3.1 Ciclo di vita di un prodotto nel commercio internazionale	17
1.3.2 Differenze tra paesi sviluppati e emergenti	18
1.3.3 Come s'identifica lo stato di sviluppo?	21
1.4 Le determinanti delle importazioni	22
2 I dati e i metodi econometrici per l'analisi empirica	25
2.1 I dati e gli algoritmi usati nello studio	25
2.2 Descrizione Dataset utilizzato	26
2.2.1 Prodotto	26
2.2.2 Popolazione	28
2.2.3 PIL	29
2.3 Descrizione delle variabili utilizzate	31
2.3.1 Introduzione	31
2.3.2 Importazioni normalizzate	31
2.3.3 Tasso di Variazione CAGR	32
2.3.4 Variazione Storica	34
2.4 Tecniche di Analisi	37

2.4.1	Machine Learning non supervisionato	37
2.4.2	Machine Learning supervisionato	38
2.5	Migliore la normalizzazione su PIL o su POP?	41
3	Analisi Empirica	43
3.1	Cluster Analysis	43
3.1.1	Risultati Analisi	47
3.1.2	Criticità analisi: gli errori di classificazione	52
3.2	Linear Discriminant Analysis	54
3.2.1	Risultati Analisi	54
3.2.2	Criticità Analisi: gli errori di classificazione	56
3.3	Implicazioni per l'esportatore italiano	59
	Conclusioni	62
	Bibliografia	63
	Appendici	65
A	Lista paesi banche dati	66
B	Comandi Python variabile Importazioni normalizzate	71
C	Comandi Python calcolo CAGR	73
D	Comandi Python Tasso Storico	75
E	Dimostrazione applicazione formule Variabile Storica	77
F	Codice Cluster Analysis	79
G	Lista Paesi Addestramento ML supervisionato	80
H	Cluster Analysis con dati normalizzati su PIL	82
I	Cluster Analysis con dati normalizzati su POP	87
J	Grafici Cluster (POP) in scala logaritmica	92
K	Grafico Cluster Crescita completo di tutti i paesi	93
L	Risultati completi LDA	94

Elenco delle figure

1.1	Fasi del ciclo vitale di un prodotto	13
1.2	Business Risk in relazione alla fase del ciclo vitale	16
1.3	Ciclo Export	18
1.4	Evoluzione importazioni/esportazioni in base alla tipologia del paese	20
2.1	Importazioni Lavastoviglie per paese 1995	27
2.2	Importazioni Lavastoviglie per paese 2021	27
2.3	Popolazione mondiale 1995	28
2.4	Popolazione mondiale 2021	29
2.5	PIL 1995	30
2.6	PIL 2021	30
2.7	Caso 1: Variazione Storica	35
2.8	Caso 3: Variazione Storica	35
2.9	Caso 2: Variazione Storica	37
2.10	Funzionamento Clustering	38
2.11	Funzionamento KNN	40
2.12	Funzionamento LDA	41
2.13	Cluster Analysis con dati normalizzati su PIL	42
3.1	Cluster Analysis normalizzata sulla popolazione ($Q/POP - tvar/POP$)	44
3.2	Cluster Analysis normalizzata sulla popolazione ($Q/POP - VarSt/POP$)	45
3.3	Cluster Analysis grafico 3D	47
3.4	Cluster paesi in Latenza senza Regno Unito	48
3.5	Cluster Paesi in Crescita con $Q/POP > 500$	49
3.6	Paesi che mostrano una Crescita molto sostenuta	50
3.7	Cluster paesi in Saturazione	52
3.8	Differenza del Regno Uniti rispetto agli altri paesi classificati come in Latenza	53
3.9	Germania, Estonia e Repubblica Ceca	54

3.10	Classificazione LDA in scala log	55
3.11	Classificazione LDA	57
3.12	Paesi Classificati in Saturazione da LDA	58
3.13	Paesi Classificati in Crescita da LDA	59
3.14	Paesi Classificati in Latenza da LDA	59
E.1	Cluster Analysis con Variazione Storica eccessivamente "pe- sante"	78
E.2	Cluster con Variazione Storica corretta	78
J.1	Cluster Analysis in scala logaritmica	92
K.1	Grafico cluster Crescita tutti i paesi	93

Introduzione

In un processo di internazionalizzazione, un fenomeno che risulta fondamentale conoscere è non soltanto il ciclo vitale del prodotto interessato (Product Life Cycle, PLC), bensì quello che potremmo definire come “ciclo vitale delle importazioni”. L’analisi delle serie storiche delle importazioni di vari prodotti in molteplici paesi rivela, infatti, come queste seguano percorsi comuni, caratterizzati dal susseguirsi di fasi temporali con caratteristiche ben distinte:

Fase 1: Latenza

Fase 2: Crescita

Fase 3: Saturazione

Individuare con accuratezza la fase di appartenenza delle importazioni in ogni paese aprirebbe alla possibilità di individuare un mercato in espansione e quindi di cogliere il miglior momento d’ingresso. Tale mossa potrebbe, nel giro di poco tempo, incrementare le esportazioni aziendali del bene di interesse verso i mercati che mostrano le maggiori potenzialità.¹

Conoscere la dinamica del ciclo vitale del prodotto rappresenta il primo punto fondamentale per riuscire a comprendere il funzionamento del ciclo vitale delle importazioni. Tuttavia, è la conoscenza di quest’ultimo che permette di cogliere importanti opportunità sui mercati internazionali. La conoscenza del ciclo vitale delle importazioni e dello stadio in cui si colloca un paese di interesse consente infatti di effettuare delle valutazioni sui tempi d’ingresso nel mercato preso in esame e individuare strategie di marketing mix più accurate e precise.²

Vi è un altro importante beneficio nel conoscere lo stadio delle importazioni e quindi di poter cogliere le migliori opportunità d’ingresso in un paese. Quando parliamo di commercio internazionale e, nello specifico, di export di un particolare prodotto, è sempre opportuno tenere in considerazione i fattori che possono influenzare la diffusione e la propensione al consumo di quel bene nel mercato di riferimento. Ad esempio, la presenza di un’infrastruttura in grado di distribuire in modo capillare il prodotto, la rilevanza di una classe

media sufficiente a sostenere la domanda dei prodotti o l'esistenza di una rete distributiva affermata e conosciuta per poter attirare nuovi clienti sono tutti imprescindibili elementi di attenzione qualora si abbia intenzione di entrare in un nuovo mercato. Quindi, riuscire a cogliere un'opportunità prima della concorrenza permette di predisporre una rete distributiva che nel lungo periodo potrà dare un vantaggio alle importazioni dei nostri prodotti. Una situazione del genere è riscontrabile sul mercato cinese per quanto concerne le importazioni di prodotti agro-alimentari, dove la presenza di punti vendita come Auchan o Carrefour fornisce un vantaggio assoluto ai prodotti francesi rispetto a quelli Italiani³.

Il presente lavoro si articola nelle seguenti parti. Il primo capitolo fornisce una panoramica sulla letteratura per individuare le caratteristiche distintive delle varie fasi del ciclo vitale delle importazioni. Il secondo capitolo descrive il processo di selezione e pulizia dei dati, con conseguente scelta delle variabili rilevanti; procede poi in una comparazione e valutazione di diversi algoritmi di Machine Learning (ML) supervisionato e non.

I dati utilizzati provengono da tre diversi dataset ottenuti grazie alla banca-dati ExportPlanning-Studiabo; il primo comprende serie storiche sulle importazioni, mentre gli altri due mostrano le serie storiche di PIL e popolazione. Gli algoritmi di ML utilizzati sono stati la Cluster Analysis per quanto riguarda il ML non-supervisionato e Linear Discriminat Analysis, Random Forest, K-Nearest Neighbors e Support Vector Machine, per il ML supervisionato. Il terzo e ultimo capitolo invece, presenta i risultati e i commenti in seguito all'applicazione dei due algoritmi che si erano dimostrati più efficienti nella fase di valutazione, la Cluster Analysis e la Linear Discriminat Analysis

Capitolo 1

Un quadro teorico di riferimento

In questo primo capitolo affronteremo, da un punto di vista storico, le diverse fasi che si susseguono nella generazione di un contesto economico e sociale che possa portare alla creazione di un mercato interno ed internazionale, descriveremo poi, il tema del ciclo vitale dei prodotti e dei rispettivi settori di appartenenza. Successivamente, basandoci su quanto descritto nelle prime due parti di questo capitolo, vedremo in che modo gli stati di sviluppo di un prodotto, influenzano le relative importazioni ed esportazioni ed infine cercheremo di fornire al lettore alcuni spunti su come l'analisi che seguirà, nei capitoli successivi, cercherà di comprendere il ciclo delle importazioni, e quindi, il momento più opportuno per investire in un determinato mercato.

1.1 Fasi di Sviluppo secondo Rostow

La teoria sviluppata da Rostow⁴ si basa essenzialmente su un approccio dinamico alla teoria della produzione. Questo perchè la teoria classica della produzione è formulata sotto rigide assunzioni statiche, che impediscono, o permettono una sola variazione, nelle variabili più importanti del processo di crescita.⁵ Così nel tempo, si è cercato di far convergere la teoria classica con l'analisi Keynesiana, arrivando così ad 'introdurre variabili dinamiche come la popolazione o la tecnologia, tuttavia lo si è fatto in maniera molto generica e mantenendo una rigidità di fondo importante, tanto da non rilevare i più importanti fenomeni di crescita.⁶

L'idea di Rostow è dunque quella di formulare una teoria dinamica della produzione in grado di individuare la composizione degli investimenti, i progressi, anche e soprattutto a livello tecnologico, in specifici settori e non solo a livello aggregato, e la distribuzione dei redditi e dei risparmi tra la popolazione di riferimento.

In relazione proprio alla dinamicità del modello, siamo in grado di individuare i punti di equilibrio di ogni settore economico e non solo in relazione a Investimenti, Consumo e Output nel complesso.⁷

Tale teoria è di fondamentale importanza anche per il proseguo di questo lavoro, poiché permette di comprendere come l'evoluzione, a livello prettamente industriale, abbia generato il contesto che noi oggi ci troviamo ad affrontare. Inoltre, come vedremo in seguito, le fasi che si susseguono nel ciclo di vita dei prodotti, e quindi del business, non sono poi così diverse, a livello concettuale naturalmente, rispetto a quelle che Rostow aveva individuato e successivamente sintetizzato per la società nel suo complesso.

Rostow dunque, individua e definisce cinque fasi di crescita che possiamo identificare come segue:

- *Società tradizionale*
- *Precondizioni per il decollo industriale*
- *Decollo Industriale*
- *Maturità*
- *Era del consumo di massa*

1.1.1 Società Tradizionale

Quando introduciamo il concetto di "Società Tradizionale", parliamo sostanzialmente di società che, pur dovendo affrontare importanti limiti legati alle proprie funzioni di produzione, si sono evolute producendo ciò che Rostow definisce come "storie di cambiamenti infiniti". Pur trattandosi di società e di realtà dotate di importanti livelli d'inventiva, nonché di grandi capacità innovative, trovano nella limitata conoscenza della tecnologia il loro limite più importante. Osserviamo infatti come più di tre quarti della forza lavoro fosse concentrata nel settore agricolo e larga parte dei consumi superiori a quelli che venivano considerati come la soglia minima per il sostentamento, veniva impiegata in spese non produttive.⁸

La centralità politica della società risiedeva nelle piccole aree, controllate dai proprietari terrieri, dove pur con tutte le tensioni tipiche dell'era esercitavano le autorità centrali dell'area.

1.1.2 Precondizioni per il decollo industriale

La Gran Bretagna è stata sicuramente il primo paese ad avviare un processo di industrializzazione costante e duraturo che ha permesso la creazione

della società odierna. È proprio con la rivoluzione industriale del 1760, che possiamo osservare come dalla "Società Tradizionale" si siano generati le pre-condizioni che hanno sostenuto la rivoluzione e che hanno portato poi alla successiva fase di crescita.

Come sostenuto da Rostow, affinché il processo di industrializzazione sia stabile ed efficiente è necessario che prima avvengano dei grandi cambiamenti in settori non industriali.

Qui di seguito possiamo trovare tali processi:

- Accumulo di capitale sociale, in particolare parliamo di ingenti "investimenti" in infrastrutture come ponti, strade o ferrovie. Questo passaggio è fondamentale in quanto permette lo sviluppo del mercato interno che porta ad un miglioramento della produttività delle risorse, consente inoltre al governo nazionale di controllare e amministrare efficientemente la nazione;
- Rivoluzione tecnologica dell'agricoltura che ne migliori l'efficienza e la produttività. L'inizio di un processo quale la rivoluzione industriale porta a un inevitabile incremento della popolazione che tenderà a concentrarsi in maniera crescente nei centri urbani, per cui se l'agricoltura non migliora la propria capacità produttiva il processo di crescita sarà inevitabilmente strozzato;
- L'aumento della produttività, delle opportunità nonché delle capacità di commercializzare offrono la possibilità di incrementare i volumi delle esportazioni, possibilità che deve essere sfruttata.

Tuttavia tali cambiamenti seguono necessariamente un'evoluzione e uno sviluppo in una "dimensione non-economica", vale a dire, per permettere uno sviluppo tecnologico nell'agricoltura, grazie al quale si migliora la produttività e l'efficienza, è necessario che la comunità agricola accetti tale cambiamento e si affidi a nuove tecniche ai più sconosciute; è necessario un governo capace non solo di mantenere l'ordine nella società, ma anche e soprattutto dotato della visione necessaria per permettere lo sviluppo di una rete di trasporti efficiente, di accumulo di capitale sociale, di una politica commerciale lungimirante e della volontà di spingere per la propagazione delle nuove tecniche agricole.

Alla ricerca di approvazione, e nel tentativo di spingere la popolazione di uno Stato verso l'accettazione dei cambiamenti a cui deve far fronte la società per affrontare il progresso tecnico, la politica cerca e cercava di esporre ai cittadini i vantaggi in termini di benessere economico, aspettativa di vita e nuove tecnologie di cui le società più avanzate dispongono. Non solo, spesso infatti

si cercava di sfruttare lo spirito nazionalistico intrinseco nella popolazione, affermando come le società più avanzate potessero in qualche modo imporre la loro volontà sugli Stati meno industrializzati. ⁹

1.1.3 Decollo Industriale

Secondo la definizione di Rostow, quando parliamo di Decollo Industriale, intendiamo il raggiungimento di una rapida crescita nel limitato gruppo di settori, che possiamo definire come settori leader, a cui sono applicate le nuove tecniche industriali.¹⁰

Rostow ritiene che anche in questa fase di crescita non debbano essere persi di vista gli sviluppi dei tre cambiamenti, processi "non industriali" individuati nelle precondizioni, egli ritiene infatti necessario proseguire il loro sviluppo allargando anche, dove possibile, gli orizzonti di sviluppo, così da permettere alla società di rimanere sul percorso di crescita ormai avviato e superare più efficacemente i vari shock che si potranno incontrare.

La teoria indica anche come, un chiaro segnale e un risultato importante della fase di Decollo Industriale risieda nella capacità della società di sostenere un tasso annuale di aumento degli investimenti netti pari almeno al 10%¹¹. Alla luce di quanto detto, quindi, possiamo affermare che una società che entra nella fase di Decollo Industriale avanzata, ossia una società che ha avviato e che sta portando avanti un importante processo di ammodernamento, sia una società che si dimostra vittoriosa, da un punto di vista politico, culturale e sociale rispetto a quelle società che hanno perseguito obiettivi che si sono rivelati errati oppure rispetto a quelle società che sono rimaste collegate alla "Società Tradizionale". Tale vittoria dovrebbe tuttavia trasformarsi in un trasferimento delle conoscenze acquisite verso le società più in ritardo nel processo di sviluppo, ciò non sempre avviene.

1.1.4 Maturità

In seguito alla fase di "Decollo Industriale" ci spostiamo in quella che viene definita come fase "Maturità". Sempre seguendo la definizione di Rostow, possiamo affermare che, una società si trova in questa fase quando sta applicando la moderna tecnologia a gran parte dei propri settori, raggiungendo così un'efficienza generale. I settori leader, ossia i settori trainanti nella fase precedente, non dipendono più esclusivamente dalla componente tecnologica che utilizzano, ma diviene invece sempre più influente la scelta a livello politico e la tipologia di risorse che quel determinato settore utilizza.

È chiaro che, oggi, in Italia, le aziende appartenenti al settore che si occupa della lavorazione del carbone non possono essere leader dello sviluppo indu-

striale italiano, anche qualora si dotassero di tecnologie all'avanguardia. Un'altra caratteristica che troviamo nelle società che si trovano a percorrere questa fase è il cambiamento della struttura e delle qualità della forza lavoro. È evidente come già dalle fasi precedenti la frazione di popolazione impiegata nel settore agricolo fosse caratterizzata da un costante decremento, decremento accompagnato da un accentramento sempre più evidente della popolazione nei centri urbani. Rispetto alle fasi precedenti, intere generazioni nascono e crescono nei centri, con una conseguente diffusa convinzione, fra questa parte di popolazione, di poter esercitare un peso maggiore sul processo politico non sentendosi più "ospite" ma parte integrante della città. Vengono così a generarsi prima, e ad aumentare poi, le richieste affinché il governo consenta e riconosca misure di supporto sociale ed economico. Un ulteriore passaggio, determinato dall'accentramento nei centri urbani è quello che porta ad un cambiamento nei vertici industriali, dove i manager della fase di Take-off, solitamente persone modeste, dotate di grande creatività che avevano avuto un'intuizione, vengono sostituiti da persone dotate di visioni molto più ampie e di maggiori conoscenze.

1.1.5 Era del consumo di massa

Come quinta e ultima fase, Rostow, individua la cosiddetta "era del consumo di massa"; in questo contesto un paese, un'economia, può essere considerata matura. La principale caratteristica di una società matura è che questa cessa di avere come principale obiettivo da perseguire l'espansione della tecnologia moderna.

Una società matura dispone di un livello di benessere sociale sufficientemente alto che le permette, sostanzialmente, di scegliere tra tre direzioni di sviluppo principali. Analizzando i Paesi sviluppati della prima metà del Novecento, possiamo osservare come ognuno di essi abbia mostrato preferenze maggiori verso una delle tre direzioni precedentemente annunciate.¹² È possibile identificare paesi che hanno concentrato gran parte dei propri investimenti, e quindi gran parte della propria attenzione politica, nel cercare di accentrare il proprio potere e la propria forza nel contesto mondiale, perseguendo quindi un obiettivo di carattere sostanzialmente militare. Altri contesti invece hanno spinto verso uno sviluppo volto al welfare, alla pubblica sicurezza e al concedere alternative al lavoro. Abbiamo, infine, paesi che hanno perseguito obiettivi legati sostanzialmente al consumo privato.

1.2 Ciclo vitale di un Prodotto

Fino ad ora abbiamo ripercorso le fasi e i processi individuati da Rostow nello sviluppo della società industrializzata nel suo complesso, di seguito affronteremo un tema che appare più contemporaneo e immediatamente applicabile al contesto del presente lavoro, ossia in che modo i prodotti e i vari settori industriali si sviluppano: il ciclo vitale di un prodotto e il conseguente ciclo vitale del settore industriale.

1.2.1 Che cos'è il ciclo vitale di un'impresa o di un settore?

Quando parliamo del ciclo vitale di un settore, di un'industria ci riferiamo a quel fenomeno che, come per gli esseri viventi, permette di suddividere l'esistenza di un'impresa in una serie di fasi aventi specifiche caratteristiche e precisi obiettivi.

Ovviamente, quando ci riferiamo al concetto di ciclo vitale lo facciamo con la consapevolezza che questo risulta essere direttamente correlato a ciò che nella letteratura viene definito Ciclo Vitale del Prodotto (PLC)

Le fasi che compongono il Ciclo Vitale ci permettono di delineare le principali caratteristiche e le principali azioni che un'impresa deve intraprendere per sostenere la crescita e il successo del prodotto o della famiglia di prodotti lanciati.

Tali fasi, che coincidono con quelle del PLC, che analizzeremo di seguito, sono:

- *Introduzione/Lancio*
- *Crescita*
- *Maturità*
- *Declino*

I modelli del ciclo vitale possono essere considerati a tutti gli effetti come modelli di Marketing, vengono infatti impiegati per descrivere e analizzare i prodotti oggetto del Business in questione, per sviluppare successivamente strategie appropriate per la specifica fase in cui ci si trova.

La Figura 1.1 mostra il grafico relativo alla curva del ciclo vitale rispetto al volume delle vendite, ai profitti generati e ai ricavi.

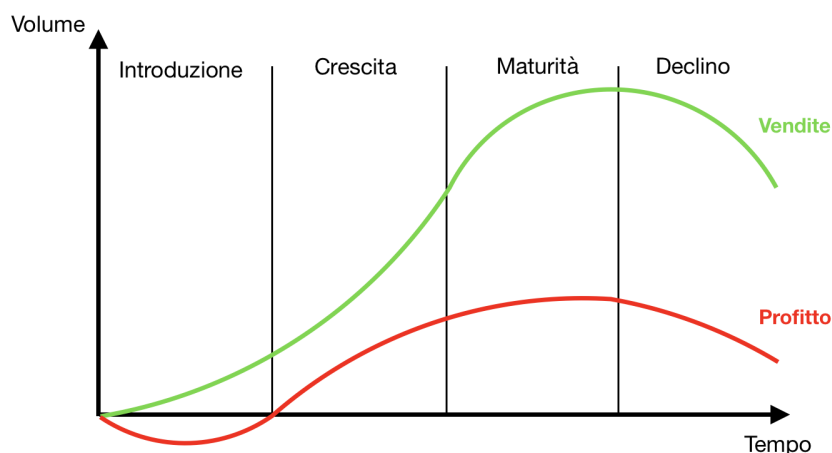


Figura 1.1: *Fasi ciclo vitale di un prodotto.* [Svend Hollensen, *Global Marketing, 8th Edition, pp.462*]

Possiamo così individuare nel volume delle vendite e in quello dei profitti le due grandezze cardine in merito alla valutazione che si può dare di un Business e alla relativa insinuazione della fase in cui si trova. Come vedremo più nel dettaglio nella prossima sezione, dove procederemo ad analizzare singolarmente le varie fasi, in Figura 1.1 vediamo come il volume delle vendite e del profitto seguano due percorsi simili ma con alcune importanti differenze. La più importante fra queste differenze è sicuramente quella relativa alla pendenza delle due curve nella fase discendente, vediamo infatti come la curva relativa al volume delle vendite abbia una pendenza decisamente superiore a quella dei profitti, questo è dovuta all'esperienza accumulata dalle imprese operanti nel settore, che, mediante ormai efficienti operazioni di scala, riescono a contenere il calo dei profitti rispetto al calo più importante delle vendite.

Risulta importante una considerazione aggiuntiva, durante la fase di declino sarebbe in linea teorica possibile presentare aggiornamenti o evoluzioni sostanziali del prodotto in grado di dare un ulteriore slancio alla curva di vendite e di conseguenze di profitto, prolungando così il ciclo vitale del Business.

1.2.2 Le quattro fasi del ciclo vitale del prodotto

Introduzione

In questa prima fase, solitamente, l'impresa promotrice del nuovo prodotto si trova isolata nell'industria. Si tratta di nuove, piccole realtà imprenditoriali

oppure di distaccamenti di realtà più importanti che sfruttano i loro reparti di ricerca e sviluppo per individuare nuovi potenziali prodotti.

Il prototipo del prodotto viene inizialmente offerto a ristrette categorie di clienti, che svolgono una funzione di stress test sul prodotto con lo specifico scopo di testarlo ed individuare i limiti, anche in relazione alla natura del prodotto. Successivamente è compito delle strategie di marketing adottate dall'impresa spiegare e generare la consapevolezza sul prodotto e sull'industria nascente. Ovviamente, in questa fase le vendite saranno ad un livello molto basso e i profitti, visti gli alti costi di marketing e di realizzazione del prodotto, saranno negativi. Per queste ragioni che solitamente gli incassi generati dalle prime vendite vengono reinvestiti nell'azienda stessa per cercare di stabilizzare la propria posizione sul mercato e proseguire nel percorso di crescita.

Spesso le aziende in questa fase cercano di raggrupparsi, anche solo a livello geografico, per rendere più facile l'accesso ai materiali chiave o a figure esperte del settore.^{13 14} L'esempio perfetto risulta essere la Silincon Valley in California.

Crescita

Anche in questa fase, come in quella precedente, è fondamentale l'apporto di capitale per sostenere il prodotto nella crescita. Il marketing del prodotto deve concentrarsi sul mostrare le differenze dell'offerta dell'impresa rispetto all'offerta dei concorrenti.

La studio in ricerca e sviluppo rimane fondamentale, poiché è probabile che sia necessario apportare modifiche al prodotto per avvicinarsi il più possibile alle vere esigenze dei clienti.

La creazione e la crescita del mercato comporteranno un inevitabile aumento della domanda e, se l'impresa sarà riuscita ad entrare nel mercato, si assisterà ad un aumento delle vendite. Come conseguenza, nel corso di questa fase, i profitti diverranno positivi e il loro tasso di crescita aumenterà.

Per quanto concerne la durata di questa fase, ciò dipende dalla tipologia del prodotto e dalla relativa industria.

Nella fase di crescita la curva del ciclo vitale risulta essere molto ripida e molto spesso in questa fase le imprese tendono a espandersi anche geograficamente, compatibilmente con la natura del prodotto.

Maturità

In questa terza fase ciò che merita di essere evidenziato è il cambio di concavità delle due curve presentate in Figura 1.1 per valutare l'andamento del

ciclo vitale; sia profitti che vendite mostrano i primi segni di quella che sarà la successiva fase di Declino.

Le vendite, cessano di crescere, si stabilizzano e si preparano ad iniziare la discesa, ammesso che non si prepari un prolungamento del ciclo vitale. Anche i profitti in seguito al calo delle vendite mostrano un iniziale appiattimento della curva, in attesa dell'inizio della discesa anche dei profitti, che tuttavia, come detto in precedenza, avverrà con un tasso di crescita minore.

La concorrenza in questa fase avviene principalmente da nuovi entranti che cercano di "strappare" quote di mercato alle imprese già presenti con prodotti solitamente di qualità inferiore, rimane quindi fondamentale per le imprese definire una strategia chiara supportata da un marketing efficace, possono infatti scegliere di proseguire nell'offerta di prodotti di qualità superiore oppure di convertirsi adottando strategie low-cost/low-price con l'intento di incrementare il volume delle vendite, sfruttando le conoscenze acquisite nelle fasi precedenti.

Un management efficiente e un buon livello di investimenti in ricerca e sviluppo dedicati al progresso del prodotto sono le migliori strategie possibili per cercare di prolungare la fase di maturità e più in generale il ciclo di vita dell'industria.¹⁵

Declino

Se l'innovazione del prodotto non è stata sufficientemente sostenuta, o se le nuove scoperte a livello tecnologico hanno fatto sì che l'industria divenisse obsoleta, il declino sarà inevitabile.

In seguito all'inizio della fase di Declino, il volume delle vendite precipiterà ad una velocità crescente, i concorrenti che non hanno ancora raggiunto la fase di Maturità con i loro prodotti usciranno dal mercato, lasciando quindi all'interno di esso solo alcune imprese che rimarranno a competere in un mercato (di nicchia) totalmente diverso dal precedente.

1.2.3 Concetto di Business Risk

Merita soffermarsi su quello che viene definito Business Risk e su come questo si evolve nel corso delle fasi che abbiamo appena descritto.

Osservando la Figura 1.2 possiamo vedere come il Business Risk, ossia il rischio d'impresa, decresca con l'avanzamento delle fasi di sviluppo dell'industria e dell'impresa. Nel momento in cui il prodotto oggetto dell'industria mostra di riscuotere successo, più aumenteranno le vendite e più l'incertezza e di conseguenza il rischio relativo a quel particolare Business diminuirà.

Sempre in Figura 1.2 è presente una fase che non abbiamo visto in precedenza, la Shake-out. Questa fase, collocandosi a metà tra la fase di Crescita e quella di Maturità evidenzia come in realtà sia in questa fase che inizia il calo del tasso di crescita delle vendite, data un'iniziale saturazione del mercato. Sempre basandosi sull'osservazione della Figura 1.2, risulta interessante notare come a fronte di un minor rischio, le possibilità di reperire capitali sul mercato siano decisamente più elevate. Le imprese mature, infatti, anche se registrano una flessione del volume delle vendite, hanno ormai raggiunto un livello di competenza, di organizzazione nonché di fiducia tali da risultare decisamente più affidabili e sicure per un investitore esterno.

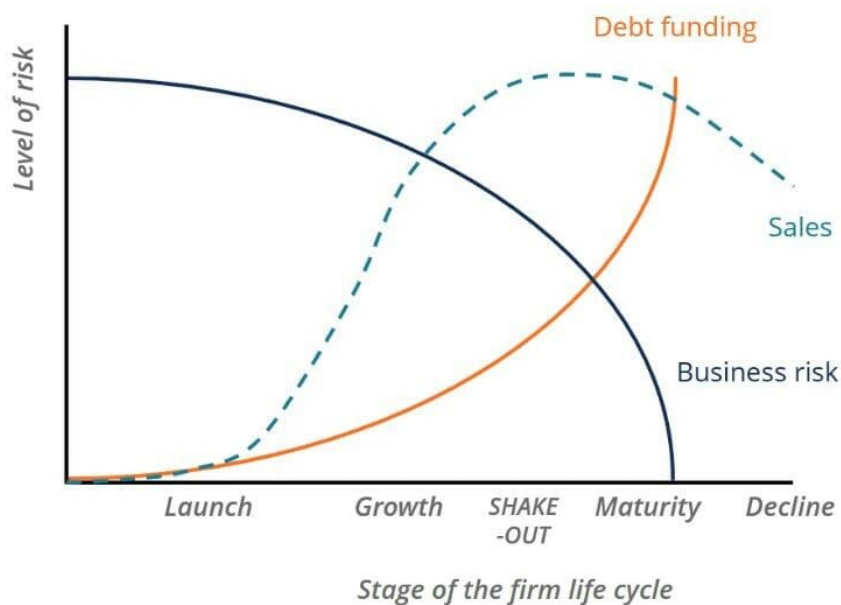


Figura 1.2: *Business Risk in relazione alla fase del ciclo vitale. [Corporatefinanceinstitute.com]*

1.3 Cicli di apertura di un paese alle importazioni di un prodotto

Come già introdotto all'inizio di questo primo capitolo, i paragrafi precedenti si sono resi necessari per esporre il contesto economico-sociale in cui gran parte degli stati, che partecipano al commercio internazionale, si trovano e le varie fasi del ciclo vitale di un prodotto oggetto del commercio.

Adesso descriveremo le modalità con cui queste fasi del ciclo vitale influenzano le relative importazioni ed esportazioni del prodotto, infine analizzeremo la modalità con cui si procederà, nell'analisi oggetto del progetto, ad una valutazione delle opportunità d'ingresso in un mercato estero.

1.3.1 Ciclo di vita di un prodotto nel commercio internazionale

Purtroppo, la letteratura si concentra principalmente sulle esportazioni, mentre il fulcro del lavoro empirico che sta dietro a questo progetto sono le importazioni. Tuttavia, a seguito di un'attenta analisi delle teorie sul ciclo vitale per il commercio internazionale è possibile individuare caratteristiche comuni che ci permettono di raccogliere qualche informazione su come trattare l'argomento di nostro interesse.

Basandoci sulla teoria di Wells¹⁶ possiamo individuare una serie di fasi in cui un determinato paese che sviluppa un determinato prodotto, passa dall'essere il principale esportatore del prodotto ad importatore con una conseguente forte concorrenza interna.

Prima di dare una spiegazione del ciclo delle importazioni credo sia necessario avere chiaro che cosa questo rappresenti.

Possiamo pensare che un determinato prodotto venga ideato e sviluppato in un determinato paese, definito come "paese natale" del prodotto. Il lancio del prodotto all'interno del paese natale permette all'azienda proprietaria di accumulare esperienza nella produzione e nella commercializzazione. A questo punto diviene probabile che, se il prodotto ha riscosso un discreto successo, inizino i primi tentativi di export. Nella fase iniziale di export, l'azienda del paese natale risulta essere praticamente monopolista nei nuovi mercati; tuttavia, se il prodotto si conferma valido e d'interesse anche al di fuori del paese natale, è probabile che le imprese estere si adoperino per avviare una fase produttiva volta alla realizzazione del nuovo prodotto. Una volta sviluppata una discreta efficienza produttiva, le produzioni delle imprese nazionali estere (rispetto al paese natale) potrebbero contare sul vantaggio dato dal risparmio dei costi di trasporto del prodotto, potendo così ampliare notevol-

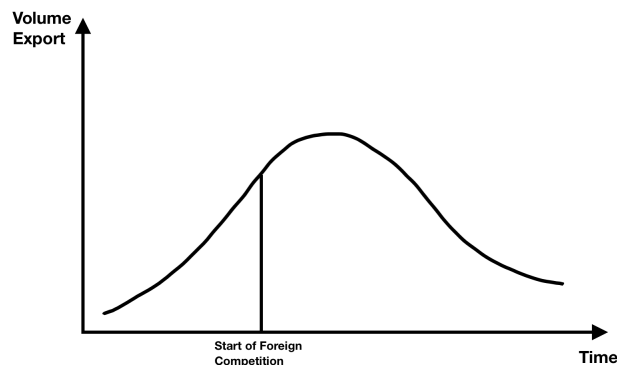


Figura 1.3: *Ciclo Export*. Louis T. Wells Jr, *A Product Life Cycle for International Trade?*, *Journal of Marketing*, Vol. 32, pp 1-6

mente la propria quota di mercato sul territorio nazionale andando così a diminuire il volume delle esportazioni da parte del "paese natale". Entriamo così nell'ultima ipotetica fase di questo ipotetico commercio, a questo punto anche il paese natale inizia un processo di importazione del bene dai paesi esteri, che ormai lo producono, comportando così un decremento delle quote di mercato anche a livello nazionale per l'azienda che aveva originariamente ideato il prodotto.

La Figura 1.3 ci mostra come il ciclo delle esportazioni può essere rappresentato.¹⁷

1.3.2 Differenze tra paesi sviluppati e emergenti

Come abbiamo visto nel paragrafo precedente le importazioni e le esportazioni seguono un determinato ciclo, dato dall'origine del prodotto e da come questo avvia il suo commercio internazionale.

Analizzando e studiando la teoria di Raymond Vernon¹⁸ vediamo come questi cicli percorrano strade differenti a seconda del livello di sviluppo del paese.

Come possiamo vedere in Figura 1.4, la prima distinzione da fare è tra "paese sviluppato ideatore del prodotto" e "altri paesi sviluppati".

Le curve Produzione e Consumo, tra le due categorie, percorrono percorsi opposti, portando a surplus di produzione, che si traducono in export, posizionati in fasi diverse relativamente al ciclo vitale del prodotto.

Secondo le più comuni teorie del commercio internazionale, le fasi del ciclo vitale sono tre e non più quattro¹, abbiamo quindi:

¹Secondo G.Lancaster e I.Wesenlund, nella loro teoria "A product Life Cycle Theory for International Trade: An Empirical Investigation", le fasi continuano ad essere quattro,

- *Nuovo prodotto*
- *Maturità*
- *Prodotto Standardizzato*

Mentre il surplus della produzione del paese ideatore, in relazione anche alla teoria di Louis T. Wells, risulta essere posizionato nella prima e in parte nella seconda fase, per gli altri paesi sviluppati risulta essere tra la seconda e la terza fase. La spiegazione a questo fenomeno ricalca perlopiù quanto detto nel precedente paragrafo.

Più attenzione merita invece il confronto tra paesi sviluppati ed emergenti. Osservando la Figura 1.4 possiamo vedere come il comportamento delle curve Produzione e Consumo abbia un andamento paragonabile a quello dei paesi sviluppati non ideatori, ma con una crescita molto più contenuta nella prima e nella seconda fase.

L'inversione tra le due curve avviene infatti in uno stato avanzato della fase di prodotto standardizzato quando ormai il prodotto è affermato a livello internazionale. Gli stati emergenti si dimostrano quindi, importatori del bene per gran parte degli stati di sviluppo del prodotto.

tuttavia mediante un'attenta analisi è evidente come Lancaster e Wasenlund spezzino in due la seconda fase, più per efficientare la spiegazione che per vera e propria necessità.

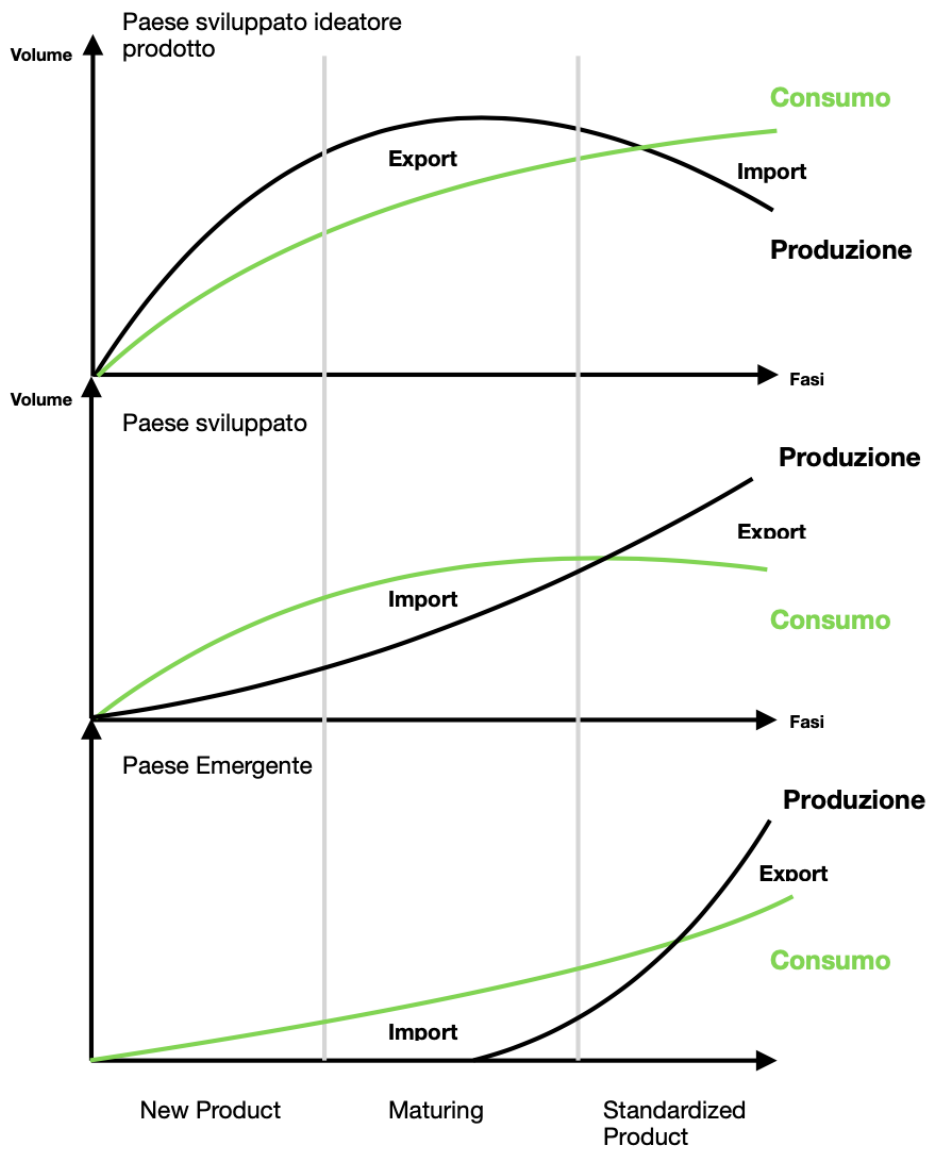


Figura 1.4: *Evoluzione importazioni/esportazioni in base alla tipologia del paese.* [Raymond Vernon, *International Investment and International Trade in the Product Cycle*, *The Quarterly Journal of Economics*, 1966, Vol.80, No.1, pp. 190-207]

1.3.3 Come s'identifica lo stato di sviluppo?

Alla luce delle considerazioni fatte sopra, ed in particolare delle teorie sul commercio internazionale^{19 20}, cominciamo ad avere gli strumenti per avanzare ipotesi su come verificare il corretto stato di sviluppo di un prodotto e di un commercio, con l'obiettivo di valutare la convenienza ad avviare un processo di esportazioni in un mercato estero.

Essendo le imprese dei paesi sviluppati più rapide ed efficienti ad avviare la produzione locale di un nuovo prodotto, i momenti migliori per un esportatore italiano sono quelli immediatamente dopo la fase di Introduzione del prodotto, quindi ipoteticamente nella sua fase di "Crescita" (par. 1.2.2 - "Crescita"). In questa fase, infatti, il consumo del bene aumenta in maniera notevole e la produzione locale è ancora debole, per cui le opportunità di profitto potrebbero essere notevoli, al netto dei costi di esportazione.

Per quanto riguarda i paesi emergenti, invece, abbiamo relativamente più margine di manovra. Come dicevamo nel precedente paragrafo, l'evoluzione delle curve Consumo - Produzione è simile quella dei paesi sviluppati ma molto più lenta nella crescita e di conseguenza, molto più lento è anche il processo che porta alla produzione locale del prodotto. In considerazione di quanto detto, quindi, le fasi migliori per avviare le esportazioni della merce in un paese emergente sono tra la fine della fase di "Crescita" e, l'inizio della fase di "Maturità". Questo perchè, per quanto comunque il consumo cresca e quindi le importazioni del bene mostrino da subito una pendenza positiva, il consumo nei paesi emergenti, soprattutto nelle prime fasi, risulta essere limitato, per cui entrare in un mercato del genere (basso consumo) con il prodotto ancora in fase di "Crescita" (fase in cui sono necessarie ancora ingenti somme di capitale per sostenere la crescita) potrebbe rivelarsi una scelta prematura che non apporta benefici in termine di volume delle vendite capaci di sostenere le spese ancora elevate, tipiche di questa fase.

Per concludere, in virtù di quanto appena descritto, risulta di fondamentale importanza per la comprensione dello stato di sviluppo delle importazioni, conoscere la differenza tra il volume delle esportazioni e delle importazioni relative allo specifico prodotto nel paese in alternativa al volume della Produzione e del Consumo. Poiché, assumendo di conoscere anche lo stato di sviluppo del paese in cui s'intende esportare, l'andamento della differenza tra esportazioni e importazioni può indicarci quali fasi del ciclo (export-import) tali paesi stiano affrontando, risultando così in grado di indicarci il momento per entrare nel mercato.

Per quanto riguarda il paese ideatore il segnale potrebbe essere nel momento in cui la differenza tra export e import, che inizialmente era crescente e positiva, comincia a mostrare i primi segnali di rallentamento nel processo

di crescita che porta ad una successiva inversione di pendenza causata dalla contrazione delle esportazioni e dal progressivo aumento delle importazioni. Per i paesi Emergenti e Sviluppati non ideatori invece, oltre alla differenza tra export e import, è di fondamentale importanza, come dicevamo all'inizio del paragrafo, la valutazione diretta sugli stati di sviluppo del prodotto.

1.4 Le determinanti delle importazioni

Nelle sezioni precedenti abbiamo affrontato il tema ipotizzando la creazione di un nuovo prodotto. Tuttavia, per fare un'analisi più accurata e utile credo sia necessaria una generalizzazione che possa così considerare tutti quei prodotti già standardizzati presenti sul mercato. Tale analisi ci permette quindi di comprendere le ragioni che impediscono ad un dato prodotto di penetrare in un dato mercato ma anche di identificare il momento più opportuno per importare il proprio prodotto in quel mercato.

Diviene così importante analizzare ed identificare le cosiddette "Barriere all'entrata di natura non tariffaria". Che cosa sono queste barriere commerciali che di fatto limitano il commercio tra stati? Possiamo identificarle in misure volontarie oppure no, che causano la limitazione delle esportazioni e delle importazioni tra Stati; possono essere considerate come barriere non tariffarie alcune regolamentazioni interne oppure le caratteristiche della rete distributiva e industriale del mercato di riferimento. In questo paragrafo vogliamo quindi offrire una panoramica su come queste "barriere" influiscono sul volume delle importazioni.

Partiamo quindi, dall'analizzare i due casi estremi:

- *Prodotto non riesce a penetrare mai un determinato mercato*: è presumibile pensare che, in determinati contesti, possano mancare le infrastrutture oppure le reti distributive sufficienti per la commercializzazione del prodotto. Si hanno quindi, una serie di vincoli di accessibilità (barriere all'entrata di natura non tariffaria) che pregiudicano l'ingresso del nuovo prodotto fino a quando i processi d'industrializzazione non avranno modificato l'assetto commerciale interno al paese.
- *Mercato nazionale saturo*: in questa circostanza invece, riprendendo il modello di Ricardo sui vantaggi comparati, abbiamo ipoteticamente raggiunto l'equilibrio tra produzione nazionale e importazioni dall'estero, per cui, teoricamente, ogni ulteriore tentativo d'ingresso da parte di altri importatori potrebbe non concretizzarsi data la saturazione (in equilibrio) del mercato.

Affrontati i casi estremi, cerchiamo ora di contestualizzare e spiegare ciò che accade nel mezzo. Negli altri casi il commercio è possibile e le opportunità concrete. Diviene così importante capire in che modo possano essere colte e quali indicatori è necessario osservare, per comprendere il momento migliore per effettuare l'ingresso nel mercato.

Come abbiamo precedentemente accennato le "Barriere all'entrata non tariffarie" possono essere di molteplici tipologie, e molto spesso i presupposti che vi stanno dietro sono totalmente diversi. L'esistenza di una normativa specifica o di certificazioni particolari sono barriere molto diverse rispetto all'assenza di una rete distributiva o di una struttura industriale. Diventa così fondamentale riuscire a capire la tipologia di barriera così da poter essere in grado di capire se queste sono superabili con alcuni accorgimenti in più sul prodotto oppure se sono insuperabili data l'assenza di una struttura alla base.

Per quanto concerne le barriere di natura legale (certificazioni o normative specifiche) possono essere superate garantendone al prodotto il rispetto; in relazione a quelle di natura sociale, quindi assenza di infrastruttura distributiva o di una struttura industriale, diviene pressoché impossibile aggirarle; infine abbiamo quelle di natura politica, frutto del tentativo dello Stato di proteggere la produzione locale da infiltrazioni estere, quest'ultimo caso è tipico dei paesi in via di sviluppo e generalmente tale categoria di barriera "affonda" con il progressivo sviluppo e apertura dello Stato al commercio.²¹ In merito alla valutazione da parte di un esportatore alla ricerca di un mercato idoneo dove investire, e alla luce di quanto appena detto, diviene fondamentale l'osservazione di due variabili in particolare:

- *Volume importazioni (standardizzato sulla popolazione)*²
- *Variazione del Volume delle importazioni negli ultimi n periodi (standardizzato sulla popolazione)*

Tali variabili, potrebbero essere in grado di spiegarci in che modo il commercio nel paese sta effettivamente evolvendo, così da fornirci le indicazioni necessarie per poter decidere se investire oppure no. A scopo puramente esemplificativo, potremmo ipotizzare 3 situazioni:

1. *Volume importazioni basso e Variazione Importazioni uguale a zero*: le importazioni sono basse e guardando agli ultimi periodi non stanno crescendo, di conseguenza è presumibile pensare che il prodotto in quello stato si trovi in una fase di latenza/introduzione;

²In base anche a ciò che l'analisi dati successiva ci mostrerà, potremmo ipotizzare dei valori soglia, sopra i quali e sotto i quali potremmo considerare il valore basso, medio o alto.

2. *Volume importazioni basso e Variazione Importazioni maggiori di zero*: le importazioni rimangono ad un livello basso, tuttavia si registra una crescita, questo potrebbe indicare un momento di transizione tra l'introduzione e la crescita (momento a livello teorico migliore per investire);
3. *Volume importazioni alto e Variazione Importazioni uguale a zero*: il paese sta importando molto, la Variazione però è tornata ad un valore prossimo allo zero, questo ci indica il raggiungimento della fase di maturità e quindi di un equilibrio nel commercio interno allo Stato tra produttori locali ed esteri.

Per concludere, le situazioni sopra riportate mostrano tre differenti scenari che fotografano il commercio di un dato prodotto in tre momenti diversi. L'obiettivo della successiva analisi sarà quindi quello di riuscire a distinguere queste fasi, cosa che ci permetterà di cogliere il momento più opportuno per investire in quel mercato.

Capitolo 2

I dati e i metodi econometrici per l'analisi empirica

2.1 I dati e gli algoritmi usati nello studio

In questo secondo capitolo, ci concentreremo sulla descrizione delle varie fasi che hanno preceduto l'analisi empirica vera e propria. Andremo a descrivere dettagliatamente i dataset che sono stati utilizzati, le variabili che è stato necessario generare ed infine presenteremo brevemente le metodologie che sta dietro agli algoritmi di Machine Learning utilizzati.

Prima di entrare nel dettaglio nei successivi paragrafi di questo secondo capitolo, ritengo sia necessaria una brevissima panoramica su quello che effettivamente troveremo.

Per quanto riguarda i dataset, ne sono stati utilizzati tre, tutti estratti dalla banca dati di Studiabo, accessibile dal sito ExportPlanning.

1. *Prodotto*: Dataset contenente la serie storica, per 151 Stati, del volume delle importazioni di Lavastoviglie, dal 1995 al 2021, espresse in milioni di \$;
2. *Popolazione*: Dataset contenente la serie storica della popolazione, dal 1995 al 2021, per 151 Stati;
3. *PIL*: Dataset contenente i valori di Prodotto Interno Lordo espresso in milioni di \$ a prezzi costanti del 2010, per 151 stati dal 1995 al 2021.

Anche le variabili, come i dataset, sono tre, ed in parte, erano state già anticipate nel precedente capitolo:

1. Q/POP oppure Q/PIL : Variabile che rappresenta il volume delle importazioni in rapporto alla popolazione o al PIL;
2. $tvar/POP$ oppure $tvar/PIL$: Variabile che rappresenta il Tasso di Variazione CAGR a 5 anni, in rapporto alla popolazione o al PIL, dove il tasso CAGR, Compounded Average Growth Rate, rappresenta la crescita percentuale media di una grandezza in un lasso di tempo²², nel nostro caso la "grandezza" sono le importazioni normalizzate e il lasso di tempo 5 anni.
3. $VarStPOP$ oppure $VarStPIL$: Variabile che rappresenta, quella che definiremo come variazione storica. Ossia, se negli ultimi 3 anni c'è stato un cambio di inclinazione nella crescita rispetto ai precedenti 10.

Infine, a conclusione del capitolo troveremo le descrizioni dei principali algoritmi di Machine Learning che sono stati affrontati:

1. *Machine Learning Non Supervisionato*: Cluster Analysis;
2. *Machine Learning Supervisionato*: Random Forest, K-Nearest Neighbors, Linear Discriminant Analysis, Support Vector Machine.

2.2 Descrizione Dataset utilizzato

2.2.1 Prodotto

Come già anticipato in precedenza, questo primo dataset contiene i dati sulle importazioni di Lavastoviglie di 151 paesi, la lista dei paesi presenti può essere consultata nella prima appendice (A).

Le Figure 2.1 e 2.2 offrono una panoramica su quelli che sono i paesi che forniscono i dati e l'evoluzione che gli ultimi 26 anni hanno portato nelle importazioni di Lavastoviglie di ciascuno di essi. Possiamo vedere come i cambiamenti più significativi si siano registrati soprattutto in Asia, dove sono molti i paesi in cui i volumi di importazioni sono arrivate a 68-147 milioni di \$, come si può osservare nella mappa di Figura 2.2 rispetto a quella di Figura 2.1. Tuttavia, il dato veramente rilevante ai fini dell'analisi sarà quello normalizzato, poiché solo questo sarà effettivamente in grado di fornirci una misura delle importazioni di ogni paese confrontabile tra stati diversi. È intuitivo capire come il dato puro, non normalizzato, non permetta un confronto diretto tra stati. È ovvio, infatti, che le importazioni ad esempio del Lussemburgo siano più basse, in livelli assoluti, rispetto a quelle della Cina; tuttavia, se andiamo a guardare il dato normalizzato la situazione si capovolge in maniera abbastanza netta.

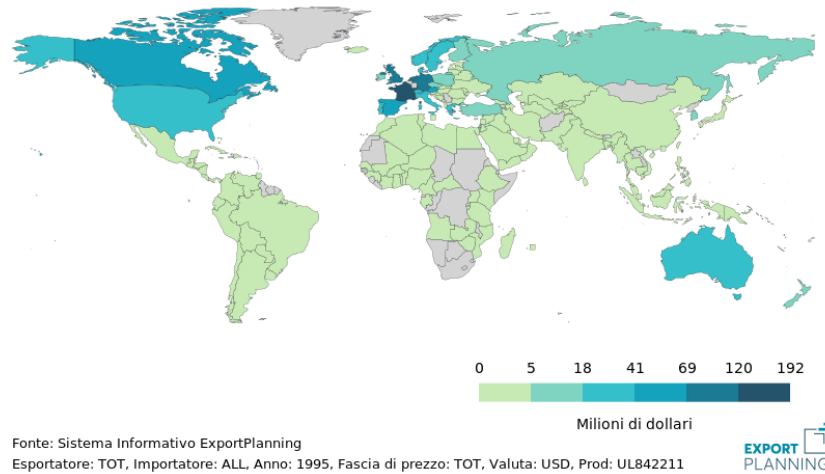


Figura 2.1: Importazioni di Lavastoviglie per paese nel 1995 (Fonte: Sistema informativo ExportPlanning)

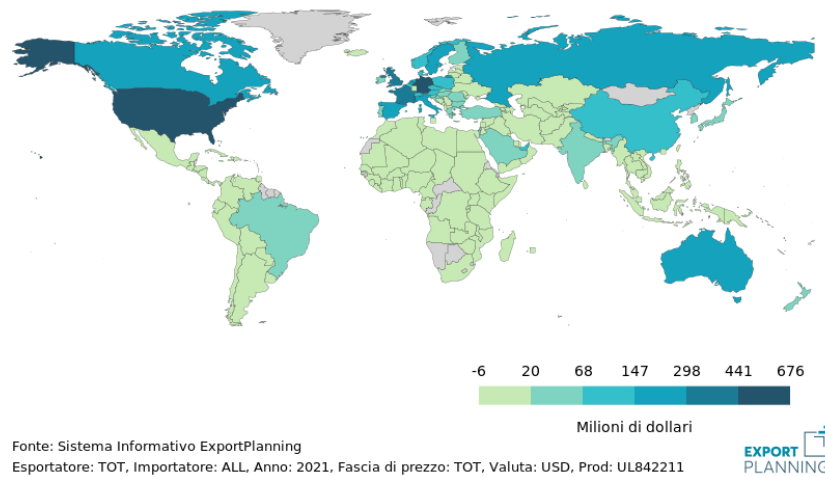


Figura 2.2: Importazioni di Lavastoviglie per paese nel 2021 (Fonte: Sistema informativo ExportPlanning)

2.2.2 Popolazione

Per quanto riguarda la popolazione in Figura 2.3 e Figura 2.4 possiamo vedere una panoramica simile a quella vista in precedenza per le importazioni. Anche in questo caso, come nel precedente, la lista totale dei paesi presenti all'interno della banca dati è stata riportata in Appendice (A). Ovviamente, non tutti i paesi che appariranno nella mappa sono stati utilizzati, poiché alcuni di questi presentavano dei dati mancanti. Su questo punto ci torneremo nel paragrafo dedicato alle specifiche variabili del lavoro. Il periodo di riferimento continua ad essere dal 1995 al 2021, in modo da coprire l'intero periodo dei dati sulle importazioni.

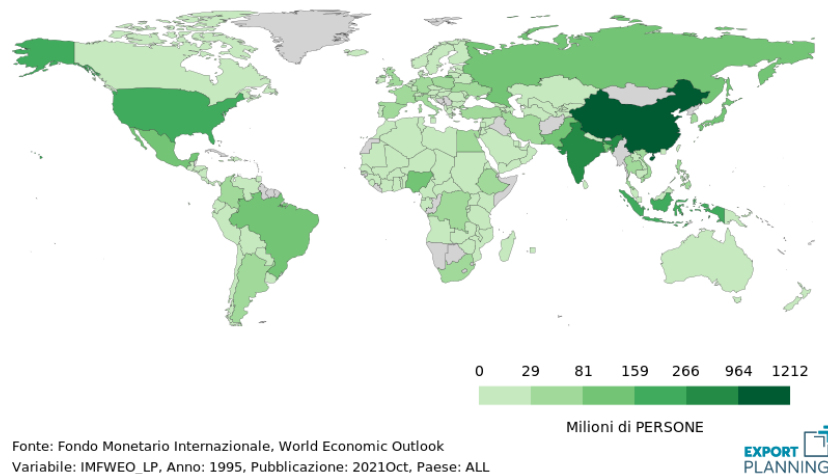


Figura 2.3: Popolazione mondiale 1995 (Fonte: FMI e WEO tramite Sistema Informativo ExportPlanning)

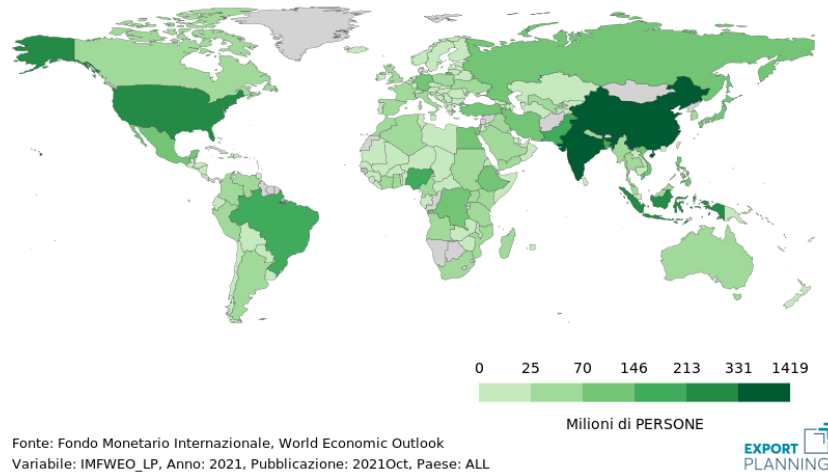


Figura 2.4: Popolazione mondiale 2021 (Fonte: FMI e WEO tramite Sistema Informativo ExportPlanning)

2.2.3 PIL

Discorso molto simile ai precedenti è anche quello che riguarda il PIL dei vari paesi. Anche in questo caso la lista completa dei paesi presenti all'interno del dataset è stata riportata in Appendice (A).

In Figura 2.5 e 2.6 possiamo vedere l'evoluzione del PIL dei vari paesi mondiali nei due periodi estremi del nostro span temporale. Come per il caso dell'import, ciò che colpisce maggiormente di queste mappe è il comparto Asiatico. È evidente come la maggiore crescita, caratterizzata dal raggiungimento di colori più scuri, sia localizzata soprattutto su Cina e India. Tale rilevanza ci permetterà considerazioni specifiche nel corso della successiva analisi empirica.

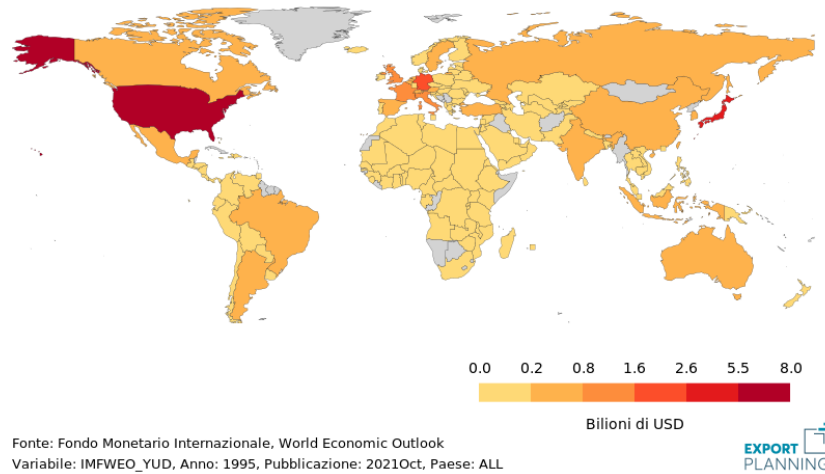


Figura 2.5: Prodotto Interno Lordo 1995 (Fonte: FMI e WEO tramite Sistema Informativo ExportPlanning)

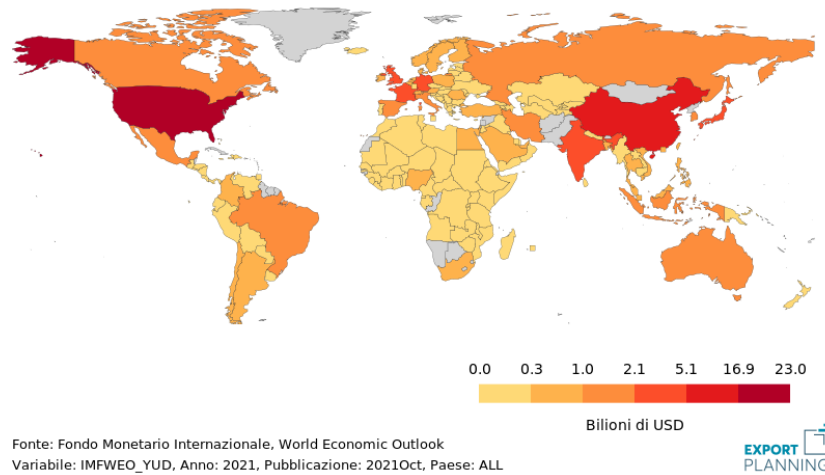


Figura 2.6: Prodotto Interno Lordo 2021 (Fonte: FMI e WEO tramite Sistema Informativo ExportPlanning)

2.3 Descrizione delle variabili utilizzate

2.3.1 Introduzione

Per procedere con l'analisi empirica è fondamentale individuare le migliori variabili da inserire all'interno degli algoritmi di Machine Learning.

Diviene quindi importante comprendere appieno la logica che sta dietro a tali variabili, in modo da poter affrontare quanto segue con una visione chiara di ciò che è stato compiuto.

Come detto nel Capitolo precedente, le tre fasi che caratterizzano le importazioni sono:

1. *Latenza*
2. *Crescita*
3. *Saturazione*

Ognuna di queste fasi presenta specifiche caratteristiche. Nella fase di *Latenza* troveremo un livello di importazioni (normalizzato) molto basso e una tendenza all'aumento contenuta. In quello di *Crescita* la variabile più indicativa sarà sicuramente il tasso di crescita, che ci si attende possa essere ai livelli più elevati. Nella fase di *Saturazione*, il tasso di crescita dovrebbe essere rientrata ad un livello in prossimità dello zero, mentre il livello delle importazioni normalizzate avrà raggiunto i valori più elevati del campione.

Prima di procedere con la presentazione delle variabili scelte, riteniamo sia importante specificare che, le variabili saranno presentate sia nella variante normalizzata sulla popolazione sia nella variante normalizzata sul PIL, ovviamente nel corso dell'analisi empirica i due percorsi saranno da considerarsi separati.

2.3.2 Importazioni normalizzate

La prima variabile che prenderemo in esame è il *Volume delle importazioni*, espresso in milioni di dollari, normalizzato sulla Popolazione o sul PIL, (Q/POP o Q/PIL).

Tale variabile, acquisisce delle potenzialità diverse in base all'abbinamento che ne facciamo; se presa singolarmente è infatti, in grado di rendere confrontabili i valori delle importazioni tra i diversi paesi, ed è in grado di fornirci un'idea, approssimativa e preliminare, sui valori soglia sopra i quali o sotto i quali si possono individuare paesi in *Latenza* e *Saturazione*. È infatti la

combinazione con le altre variabili che ci consentirà una più accurata classificazione.

La combinazione con il tasso di variazione CAGR, ad esempio, ci permette di classificare con maggiore accuratezza i vari paesi quando sono stabilmente all'interno di una fase e inoltre è in grado di individuare quelli che si trovano nei momenti finali dello stadio di *Crescita*. Infatti, se uno Stato dispone di un volume di importazioni elevato e un tasso CAGR contenuto, è molto probabile che abbia raggiunto la Saturazione, mentre se il volume è basso ma il CAGR è elevato potremmo trovarci di fronte ad uno stato in fase di Crescita.

Modalità di calcolo

In questa sezione, spiegheremo a parole in che modo la variabile è stata ricavata. Per quanto riguarda il codice Python relativo, rimando all'apposita sezione dell'Appendice (B).

Una volta normalizzati tutti i dati relativi alla serie storica delle importazioni, con il valore della popolazione o del PIL, abbiamo calcolato la media aritmetica delle osservazioni del 2018, 2019 e 2020. In un primo tentativo era stato inserito anche il dato 2021 al posto di quello del 2018, tuttavia la presenza di due anni 2020 e 2021 influenzati dal Covid, distorceva in maniera eccessiva il valore, restituendo, di fatto, risultati dalla difficile interpretazione.

In generale, la selezione delle medie su tre anni è giustificata da alcuni fattori: il primo riguarda i ritardi dovuti alla pubblicazione dei dati per alcuni paesi, che tendono ad accumulare ritardi, per cui possono capitare anni in cui quasi non risultano importazioni mentre l'anno successivo risultano livelli decisamente superiori a quella che era la media precedente. Il secondo motivo riguarda, il tentativo di smussare le oscillazioni di breve periodo, che possono essere legate a molteplici fattori, climatici, geopolitici, sanitari, che fanno sì che l'osservazione puntuale di un anno non sia sufficientemente affidabile.

2.3.3 Tasso di Variazione CAGR

La seconda variabili che abbiamo individuato è il *Tasso di variazione* ($tvar/POP$ o $tvar/PIL$), tale indicatore ci permette di comprendere l'andamento delle importazioni normalizzate negli ultimi 5 anni, fornendoci quindi una parte delle informazioni necessarie ad una corretta classificazione. Come precedentemente annunciato, per sintetizzare questo dato abbiamo deciso di utilizzare un tasso CAGR.

Poiché il nome $tvar/POP$ e $tvar/PIL$ potrebbe trarre in inganno riteniamo sia necessaria una piccola spiegazione. Tali nomi non si riferiscono ad una

normalizzazione del tasso di variazione, bensì indicano che il valore delle importazioni su cui si è costruito il tasso di variazione è normalizzato sulla popolazione o sul PIL.

Modalità di calcolo

Come prima, a questa descrizione sommaria corrisponde una procedura in Python, riportata in Appendice (C)

Nella formula 2.1 possiamo osservare l'equazione su cui si basa questo tasso di variazione:

$$CAGR = \left(\frac{EV}{BV}\right)^{1/n} - 1 \quad (2.1)$$

dove :

EV = Valore Finale

BV = Valore Iniziale

n = numero di periodi

Nel nostro caso specifico quindi, il CAGR assumerà la seguente forma:

$$CAGR = \left(\frac{Q^{norm}_{2021}}{Q^{norm}_{2016}}\right)^{1/5} - 1 \quad (2.2)$$

Per quanto concerne il numero di periodi del CAGR, la scelta è ricaduta sui 5 anni, poiché è la una lunghezza che ci permette una visione più dettagliata degli ultimi anni, senza un'eccessiva influenza passata, che avrebbe causato inevitabilmente una sovrastima su gran parte delle osservazioni, in particolare su quei paesi che hanno raggiunto o stanno raggiungendo lo stadio di Saturazione.

Essendo la tendenza di fondo, di quasi tutti i paesi, e in particolare dei paesi più sviluppati (cioè coloro che hanno raggiunto prima l'ultimo stadio) crescente, l'inclusione di un periodo eccessivamente lungo (ad esempio 10 anni) avrebbe rischiato di attribuire un CAGR rilevante anche a paesi che ormai sono nella fase di Saturazione da tempo, ma che a seguito degli effetti dal CAGR a 10 anni sarebbero stati classificati nello stadio di Crescita.

2.3.4 Variazione Storica

L'ultima variabile che affronteremo è quella che abbiamo definito come *Variazione Storica* ($VarSt/POP$ e $VarST/PIL$). Questa è stata senza dubbio la più complicata da immaginare e sintetizzare. Lo scopo principale di tale variabile è quello di rilevare la presenza di un cambio di inclinazione nella crescita delle importazioni normalizzate. Un cambio di inclinazione positivo è, infatti, una delle caratteristiche più distintive dell'ingresso nella fase centrale dello stadio di *Crescita*.

Modalità di calcolo

Come prima, in Appendice (D) si trovano i codici Python per la costruzione di quest'ultima variabile.

Per il calcolo di questa variabile si è reso necessario pensare alle diverse situazioni che si sarebbe potuto incontrare e procedere quindi di conseguenza. In Figura 2.7, 2.8 e 2.9, possiamo vedere tre distinte situazioni che hanno comportato metodologie di calcolo differenti.

Innanzitutto, procediamo con il definire le variabili necessarie alla costruzione della *Variazione Storica*.

- $M0$: Rappresenta una media pesata delle importazioni normalizzate del 2020 e del 2019. (distribuzione dei pesi: 55% il 2020 e 45% il 2019);
- $M3$: Rappresenta la media aritmetica delle importazioni normalizzate inerente agli anni 2016, 2017 e 2018;
- $M10$: Rappresenta la media aritmetica delle importazioni normalizzate inerente agli anni 2006, 2007 e 2008;
- $M'10$: Valore teorico di $M10$, assumendo una crescita costante a partire dalla 1995 fino al valore rilevato in $M3$;
- $D3$: Differenza tra $M0$ e $M3$;
- $D10$: Differenza tra $M3$ e $M10$;
- $DM10$: Differenza tra $M'10$ e $M3$.

Per quanto riguarda il calcolo, possiamo procedere analizzando le formule sotto riportate:

$$VarSt = \frac{D3}{D10} \text{ se } D10 \geq DM10 \quad (2.3)$$

$$VarSt = \frac{D3}{DM10} \text{ se } D10 < DM10 \quad (2.4)$$

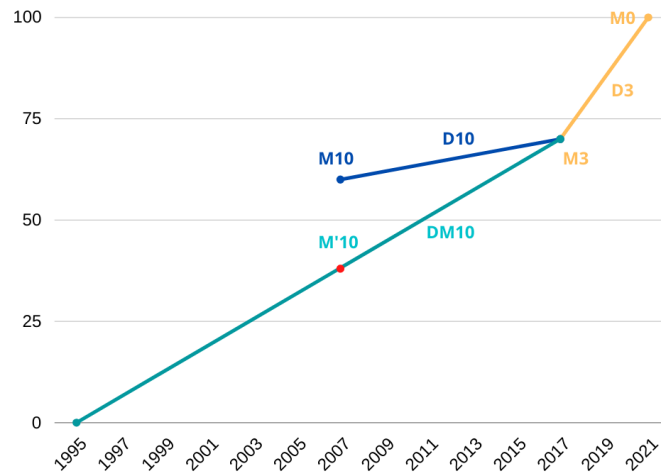


Figura 2.7: Caso N.1 di calcolo Variazione Storica. *Fonte: Elaborazione personale*

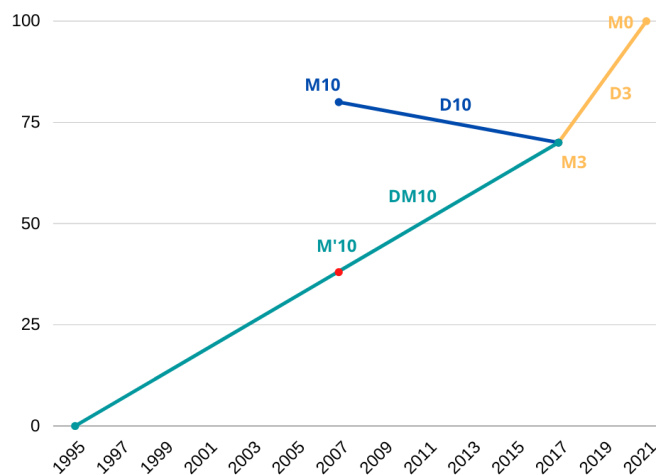


Figura 2.8: Caso N.2 di calcolo Variazione Storica. *Fonte: Elaborazione personale*

La Formula (2.3), è associata alla situazione di Figura 2.7 e 2.8, mentre la Formula (2.4) si applica al caso di Figura 2.9.

La creazione di due formule si è resa necessaria, per risolvere una situazione come quella in Figura 2.8, dove $D10$ assumendo un valore negativo, una volta

inserito nella Formula (2.3) non avrebbe dato alcun risultato interpretabile nella nostra analisi. Una possibile soluzione sarebbe potuta essere quella di prendere i valori assoluti della variazioni, ma facendo così non avremmo ottenuto l'informazione che cercavamo dalla variabile in questione, ossia non sarebbe stata in grado di evidenziarci l'entità del cambio di inclinazione e quindi avremmo rischiato di perdere una parte di tutti quei paesi che dallo stato di Latenza stanno entrando in quello di Crescita³.

Di conseguenza, per ovviare al problema dei valori negativi si è optato per ipotizzare una crescita costante, da zero a $M3$. Per ottenere quindi il valore $DM10$ abbiamo svolto il seguente calcolo: $DM10 = (M3/n) * p$ dove $n = 25$ rappresenta i periodi totali in cui si dispone di osservazioni (1995-2020) e $p = 10$ rappresenta il numero di anni nel periodo $M10 - M3$.

Per quanto riguarda il caso della Figura 2.7, esso giustifica perchè si è scelto di estendere l'utilizzo della Formula (2.3) non solo agli stati che presentavano valori di $D10$ negativi, ma anche a quelli che hanno registrato tra il 2007 e il 2017 una crescita minore rispetto a quella teorica ($DM10$). Imponendo una condizione meno restrittiva, che quindi limitasse l'applicazione della Formula (2.3) ai soli paesi con $D10 < 0$ si andava ad avere valori della variabile eccessivamente "pesanti" rispetto alle altre variabili, andando così a compromettere l'analisi Cluster. In Appendice (E) possono essere consultate le diverse prove che hanno portato a tale decisione.

In Figura 2.9 possiamo osservare invece la situazione in cui si procede con l'applicazione della Formula 2.4. Tale situazione è tipica di quei paesi che hanno da poco iniziato la fase di Crescita sostenuta, come ad esempio gli Emirati Arabi Uniti, caratterizzati da valori della *Variazione Storica* molto vicini all'unità. Unità che ci indicherebbe come negli ultimi 3 anni si sono avute delle importazioni normalizzate uguali ai precedenti 10.

Ovviamente per i paesi in fase di Latenza e Saturazione, in cui il più delle volte si va ad applicare la Formula 2.3, i valori che si ottengono sono ben al sotto dell'unità, trovandosi spesso intorno allo zero.

³La *Variazione Storica* ha, come unico scopo, quello di indicare se negli ultimi 3 anni si è registrato un incremento rilevante rispetto ai 10 anni precedenti. Prendere il valore assoluto di $D10$, voleva dire rischiare di non riuscire ad individuare tutti quei cambi di inclinazioni successivi ad un momento di decrescita.

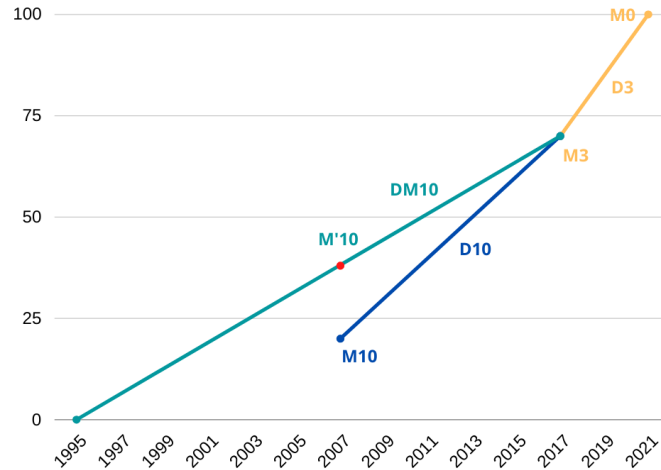


Figura 2.9: Caso N.3 di calcolo Variazione Storica. *Fonte:Elaborazione personale*

2.4 Tecniche di Analisi

In quest'ultima sezione di questo secondo capitolo, ci occuperemo di descrivere quelle che sono state le tecniche di Machine Learning (ML) che sono state utilizzate per lo svolgimento dell'analisi oggetto di questo lavoro. Inoltre, verranno descritte le scelte che sono state fatte ai fini dell'analisi.

Come precedentemente anticipato, si è scelto di intraprendere due strade differenti, una riguardante l'utilizzo di tecniche di Machine Learning non supervisionato ed in particolare la Cluster Analysis; l'altra riguardante le tecniche di ML supervisionato come Random Forest, K-Nearest Neighbors, Linear Discriminant Analysis e Support Vector Machine.

2.4.1 Machine Learning non supervisionato

Quando parliamo di Machine Learning (ML) non supervisionato, ci riferiamo ad una classe di tecniche di "classificazione" in cui abbiamo a che fare con dati senza etichetta oppure dati la cui struttura è ignota.²³

Il Clustering è una delle più importanti tecniche non supervisionate che ci consente di organizzare le informazioni in sottogruppi significativi (i Cluster), senza disporre originariamente di alcuna appartenenza a tali gruppi. Non è complicato da comprendere come, nel corso della nostra analisi, l'obiettivo è stato quello di individuare 3 gruppi distinti, dove ognuno di essi fosse in grado di rappresentare uno degli stadi di sviluppo delle importazioni (*Latenza*,

Crescita e Saturazione).

Ai fini computazionali l'algoritmo che si è scelto di utilizzare è il *k-means*⁴. L'algoritmo *k-means* fa parte della classe di Clusternig basati su prototipi. La caratteristica principale di questa categoria è che ogni cluster individuato è caratterizzato da un prototipo²⁴, che può essere identificato in *centroide* (la media dei punti simili), o in *metoide*, (il punto più rappresentativo del cluster). La ragione principale per cui si è scelto di utilizzare tale algoritmo, tuttavia, ricade proprio in quella che viene considerata come la sua più grande criticità, il più importante problema, l'individuazione del numero *k* di Cluster. Se da un lato, è in grado di individuare in maniera efficiente i Cluster nei dati, dall'altro è necessario fornirgli a priori il numero *k* di Cluster da individuare, nel nostro caso, interessandoci gli stati di sviluppo delle importazioni che abbiamo visto essere tre, è stato relativamente semplice individuare tale valore di *k*. In Figura 2.10 possiamo osservare in che modo l'algoritmo opera.

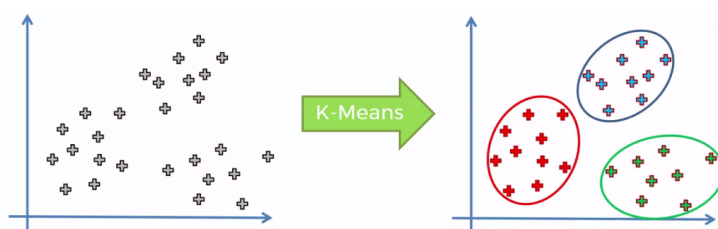


Figura 2.10: Meccanismo di funzionamento del Clustering con k-means (Fonte: <https://medium.datadriveninvestor.com/k-means-clustering-ac3ff1d3463d>)

2.4.2 Machine Learning supervisionato

Con Machine Learning supervisionato facciamo riferimento a quelle categorie di tecniche che permettono di istruire un modello, mediante la fornitura come input di una serie di dati già etichettati, che permettono così di eseguire predizioni o classificazioni di dati diversi da quelli di apprendimento.

Nel nostro caso, abbiamo assegnato a 24 paesi (Appendice G) l'etichetta rappresentante lo stadio di sviluppo delle importazioni che riteniamo più corretta. Ovviamente tale decisione è stata presa tenendo in considerazione i risultati ottenuti nell'esecuzione della Cluster Analysis. Sempre in appendice (G) possono essere consultati anche i codici inerenti tale metodologia.

Qui di seguito possiamo vedere la Tabella 2.1 in cui sono riportati i risultati ottenuti da alcuni test di accuratezza che sono stati svolti con i diversi

⁴Come per i paragrafi precedenti, in appendice (F) si può consultare il relativo codice Python.

Tabella 2.1: Risultati dei test sui vari algoritmi *ML non supervisionato*

	Training Set	Test Set
Random Forest	1.00	1.00
K-Nearest Neighbors	0.93	1.00
Linear Discriminant Analysis	0.93	0.90
Support Vector Machine	0.97	0.80

algoritmi, test che si sono resi necessari per poter successivamente scegliere l'algoritmo migliore per l'analisi vera e propria che seguirà nel prossimo capitolo. Il test è stato svolto individuando in maniera casuale un sotto campione dei dati (tutti già classificati), i quali sono stati divisi in due gruppi, anch'essi casuali, dove il primo gruppo rappresentava i paesi necessari per l'apprendimento, il secondo quelli su sui si svolgeva il test. I risultati del test sono espressi in una scala che va da 0.0 a 1.0 , dove 1.0 indica che tutti i paesi sono stati classificati correttamente.

Random Forest

La logica di fondo sul funzionamento di questa tipologia di algoritmi è quella di unire un gruppo di modelli ritenuti deboli (gli alberi decisionali) in modo da crearne uno forte e aggregato. La principale criticità di questi modelli risiede nella loro troppo elevata capacità di adattamento, che permette buoni risultati nella fase di addestramento (si osservi la Tabella 1) dove risulta essere addirittura il modello migliore, ma che ha una performance meno soddisfacente con i dati reali. Esplicitamente: un eccesso di descrizione dei dati (overfitting) porta a predizioni poco affidabili perchè il modello trattiene troppo rumore e così "sente" peggio il segnale. Anche nel nostro caso, infatti, la scelta finale non è ricaduta su questa metodologia che aveva offerto risultati migliori.

K-Nearest Neighbor (KNN)

Il K-Nearest Neighbor viene comunemente definito come algoritmo "pigro", poiché non vi è un apprendimento di una funzione discriminativa in grado di comprendere la logica dietro i dati, bensì avviene una memorizzazione del dataset di addestramento.²⁵ A questo punto la procedura prevede di verificare i k più prossimi elementi dell'addestramento (nel nostro caso $k=5$), verrà eseguita una semplice media tra questi elementi, ed in seguito quindi, classificato in conseguenza alla media ottenuta il dato sotto analisi.

In Figura 2.11 possiamo osservare come opera questo algoritmo. È immediato notare come, il "Dato da Classificare", al termine della valutazione verrà classificato come appartenente al gruppo "Caratteristica 1", poiché nei k più prossimi elementi ben 3 appartengono anche questo gruppo.

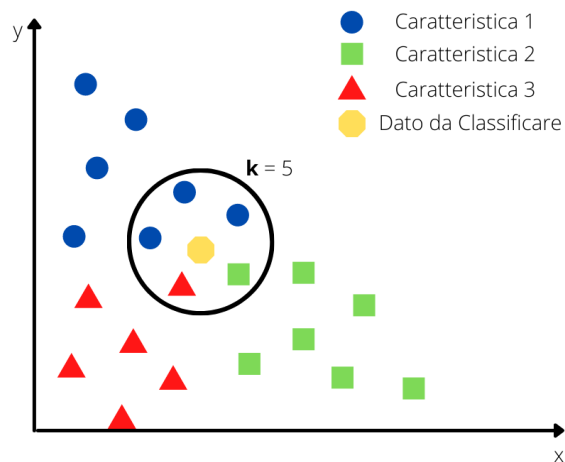


Figura 2.11: Funzionamento Classificatore KNN [Fonte: Rielaborazione Personale]

Support Vector Machine (SVM)

Gli algoritmi di SVM nascono per risolvere i problemi di classificazione non lineare. L'algoritmo, infatti, riesce ad individuare un iperpiano in grado di dividere il set di dati nel numero di caratteristiche presenti.²⁶ Nella nostra situazione SVM, l'iperpiano sarà un piano in tre dimensioni, una per ognuna della fase di sviluppo. La principale criticità di questa famiglia di algoritmi è l'essere data intensive riguardo alla dimensione del Training Set, che nel nostro caso non è sufficientemente ampio per offrire una classificazione soddisfacente.

Linear Discriminant Analysis (LDA)

Questa tecnica è una procedura molto simile alla Principal Component Analysis. Ma, mentre quest'ultima ha l'obiettivo di individuare gli assi dei componenti ortogonali che possano spiegare maggiormente la varianza del dataset, la LDA ha l'obiettivo di individuare il sottospazio delle caratteristiche in grado di ottimizzare la separabilità delle classi.²⁷ Essendo le nostre classi, in un quantitativo relativamente contenuto, tale algoritmo è stato in grado

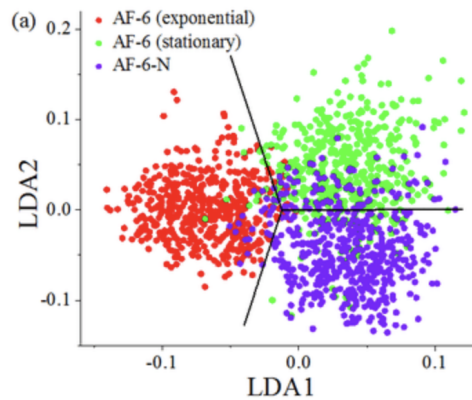


Figura 2.12: Funzionamento Linear Discriminant Analysis [Fonte: https://www.researchgate.net/figure/Results-of-the-LDA-analysis-and-Students-t-test-to-show-the-statistical-differences_fig_5304630657]

di apprendere e classificare poi i dati nella maniera più efficiente, tra gli algoritmi supervisionati.

In Figura 2.12 possiamo osservare un esempio, di come avviene la divisione mediante LDA. Possiamo vedere infatti come, quello che l'algoritmo cerca di fare è di suddividere in maniera ottimizzata il piano, in modo da garantire il più possibile una separabilità delle classi, così da riuscire a classificare il data analizzato in base alla posizione che assume nel piano.

Essendosi dimostrato quello più affidabile, la scelta finale su quale metodologia utilizzare nel corso dell'analisi è ricaduta su quest'ultimo metodo.

2.5 Migliore la normalizzazione su PIL o su POP?

Arrivati a questo punto, dopo aver presentato tutti i dati che sono stati utilizzati e tutte le variabili che si sono rese necessarie per l'analisi, diviene fondamentale indicare e contestualizzare i motivi per cui, nel prossimo capitolo verrà presentata solamente l'analisi relativa ai dati normalizzati sulla popolazione.

Come abbiamo indicato nel corso di tutto questo capitolo, si è scelto di percorrere, ai fini della normalizzazione, due strade parallele, la prima utilizzando la Popolazione la seconda utilizzando il PIL.

Arrivati però all'analisi vera e propria, e in particolare già dopo aver svolto la Cluster Analysis, si è visto che nel percorso relativo al PIL, c'era qualcosa che non andava. In Figura 2.13 possiamo osservare come il Cluster 0 (Satu-

razione) sia assolutamente sovrappopolato⁵.

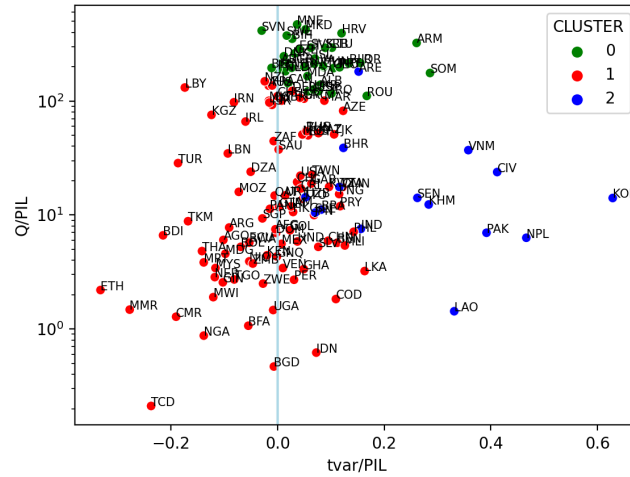


Figura 2.13: Cluster Analysis con dati normalizzati su PIL

A seguito della consultazione dell'appendice si può notare come, il problema principale che risiede dietro questo percorso, è probabilmente dovuto ad un livellamento eccessivo nella variabile Q/PIL , la quale mostra un range di valori (0-400) decisamente più contenuto rispetto, ad esempio, a quello Q/POP ⁶(0-20000). Tale range, impedisce una suddivisione efficiente dei Cluster e come conseguenza finale abbiamo che il gruppo dei paesi in Crescita (Cluster 2 di Figura 2.13) praticamente non risulta rappresentato.

Vi è inoltre un'altra problematica che genera questa sottorappresentazione degli stati in Crescita, la soglia di $tvar/PIL$ sotto o sopra la quale si può individuare uno stato nello stadio di Crescita è eccessivamente spostata a destra (circa 0,15), conseguentemente questo porta anche ad un sovrarappresentazione degli stati in Latenza. Tale problematica potrebbe essere individuata nella distribuzione della variabile $tvar/PIL$, che a sua volta era stata generata dall'utilizzo della formula CAGR sulle importazioni normalizzate.

In seguito a queste problematiche, che distorcevano eccessivamente i vari Cluster, si è quindi deciso di proseguire solamente con l'analisi mediante dati normalizzati esclusivamente sulla popolazione.

⁵In Appendice (H) può essere consultata la tabella con tutti i paesi divisi per Cluster

⁶In Appendice (I) si può consultare le tabelle con i Cluster relativi alle variabili normalizzate sulla popolazione

Capitolo 3

Analisi Empirica

Nel corso di questo capitolo, si introdurranno i principali risultati che abbiamo ottenuto dall'analisi, i cui risvolti si concentrano sull'indicazione di quali sono le opportunità d'investimento per un potenziale esportatore. Essendo l'obiettivo del progetto quello di classificare i vari stadi di sviluppo delle importazioni in diversi paesi, una corretta classificazione di uno stadio di "crescita" o di "latenza", permette di poter effettuare valutazioni più accurate nell'ottica di un investimento, rispettivamente, nel breve e nel medio-lungo termine.

Per sintesi, sono stati riportati solamente i risultati dei due algoritmi che si sono dimostrati maggiormente efficienti, la *Cluster Analysis*, unico algoritmo di *ML non supervisionato* che abbiamo introdotto nel capitolo precedente, e la *Linear Discriminant Analysis (LDA)* per quanto riguarda gli algoritmi di *ML supervisionato*.

L'ordine in cui gli algoritmi sono presentati è parte integrante del progetto, poiché, i risultati che si sono ottenuti nella *Cluster Analysis* hanno fornito di fatto, i vari paesi "soglia" necessari per l'apprendimento della LDA.

Come forma di presentazione dei risultati si è scelto di offrire una panoramica sulle evidenze che ognuno degli algoritmi è stato in grado di fornire, in seguito, una volta individuati gli errori di classificazione, si è cercato di commentarli e contestualizzarli in modo da comprenderne la natura.

3.1 Cluster Analysis

Come descritto nel capitolo precedente, l'algoritmo di Clustering che abbiamo utilizzato è il *k-means*.

Sulla base delle considerazioni teoriche espresse nel primo capitolo, ciò che ci attendiamo di osservare, grazie alla analisi di seguito proposta, è un di-

verso raggruppamento dei paesi che si collocano nelle tre fasi di sviluppo rispetto alle tre variabili prese in considerazione: importazioni pro-capite normalizzate (Q/POP), tasso di variazione ($tvar/POP$) e variazione storica ($VarSt_POP$). In particolare, ci aspettiamo che i paesi in stato di saturazione risultino caratterizzati da un livello di importazioni elevato associato a livelli prossimi allo zero per le variabili variazione storica e tasso di variazione; per quanto attiene ai paesi che si collocano in uno stato di latenza, prevediamo di osservare valori delle importazioni particolarmente bassi congiuntamente a tassi di crescita e variazione storica anche negativi. Infine, ci attendiamo che i paesi in stato di crescita si collochino in una situazione di media per quanto afferisce alle importazioni e tassi di crescita e variazione storica elevati.

Le Figure 3.1 e 3.2 riportano i risultati definitivi della classificazione e ci consentono di operare un confronto fra quanto ci aspettavamo di osservare e ciò che realmente osserviamo.

La Figura 3.1 permette di apprezzare maggiormente in che modo è avvenuta la classificazione rispetto a Q/POP e $tvar/POP$ (sugli assi), mentre in Figura 3.2 vediamo la stessa classificazione, ma rispetto ancora a Q/POP sull'asse verticale e alla Variazione Storica ($VarSt_POP$)⁷ sull'asse orizzontale.

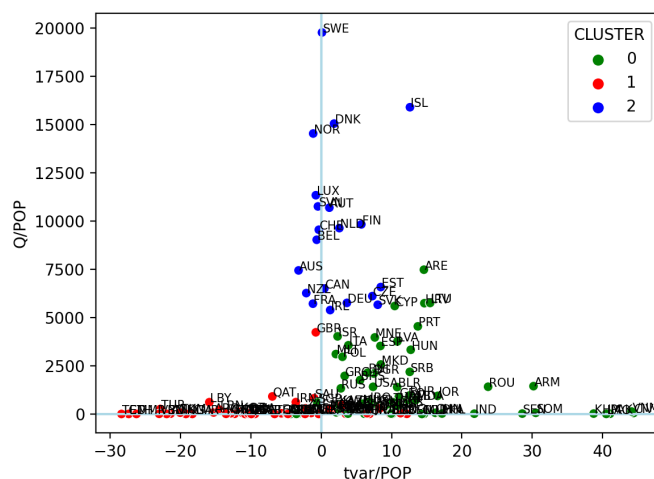


Figura 3.1: Cluster Analysis normalizzata sulla popolazione (Q/POP - $tvar/POP$)

Dall'osservazione di queste figure possiamo trarre una serie di considerazioni che ci aiutano a capire ciò che vi è dietro a quello che stiamo effettiva-

⁷In Appendice (J) si riportano gli stessi grafici ma in scala logaritmica, in modo da apprezzare maggiormente i tanti paesi che si trovano con valori di Q/POP vicini a zero.

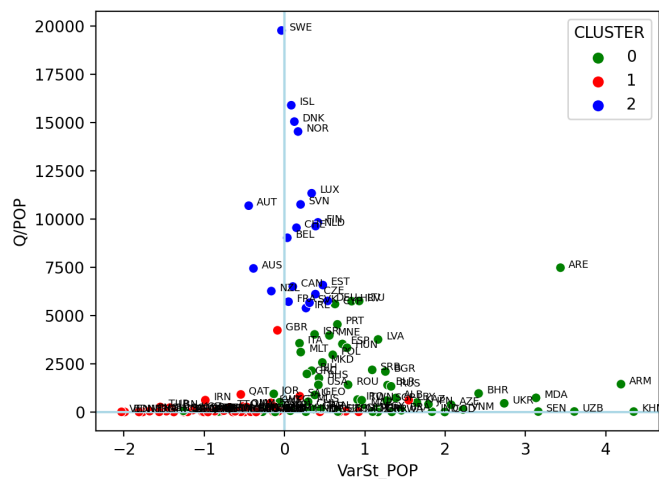


Figura 3.2: Cluster Analysis normalizzata sulla popolazione (Q/POP - $VarSt/POP$)

mente osservando.

I paesi nello stadio di Saturazione, come possiamo vedere dalla Figura 3.1, sono distribuiti esattamente dove ci aspettavamo, attorno a $tvar/POP=0$ e ad un livello decisamente più alto (in media) agli altri gruppi, rispetto alle *importazioni normalizzate* (Q/POP). Possiamo avere conferma anche dai dati della Tabella 3.1. I paesi in Saturazione, cioè i paesi dove le *importazioni normalizzate* hanno raggiunto un livello tale per cui hanno smesso di crescere, dispongono del valore medio maggiore nella variabile Q/POP , il che è coerente con le aspettative, di una variazione $CAGR^8(tvar/POP)$ media molto contenuta e una *Variazione Storica* ($VarSt_POP$) tendente a zero, tutte caratteristiche che inquadrano perfettamente tale stadio di sviluppo, dove si cresce poco poiché il livello raggiunto ormai è tendenzialmente molto alto e il mercato è saturo. Anche i valori medi del *Tasso di variazione* degli altri Cluster ricalcano perfettamente quello che ci saremmo aspettati, dove il valore più alto è riconducibile ai paesi in Crescita e quello più basso a quelli in Latenza. Infine, abbiamo la *Variazione Storica*, anche qua registriamo un'altra conferma, nei paesi in Crescita siamo ad un valore vicino all'unità, valore soglia di cui avevamo parlato anche nel precedente capitolo, mentre per i paesi in Latenza, il valore medio della variabile si assesta addirittura ad un livello minore di zero, indicandoci che, le importazioni negli ultimi 3 anni

⁸Compounded Average Growth Rate, rappresenta la crescita percentuale media di una grandezza in un lasso di tempo

oltre ad essere state inferiori rispetto ai 10 anni precedenti stanno registrando anche una tendenza al ribasso.

Tabella 3.1: Valori medi delle variabili nei Cluster

Cluster	N. Paesi	Q/POP	tvar/POP	VarSt_POP
Latenza (Cluster 1)	52	197.87	-7.19	-0.65
Crescita (Cluster 2)	67	1310.52	11.79	0.99
Saturazione (Cluster 0)	20	9567.34	2.17	0.15

Dall'osservazione del grafico in Figura 3.3 possiamo vedere da varie angolazioni, il movimento che i vari stati fanno nel passare da uno stadio ad un altro. Si vede infatti, come il Cluster 2 (Crescita) si espanda molto di più rispetto agli altri, per poi però convergere con l'aumentare delle importazioni pro-capite (Q/POP) nella zona in cui sono i paesi in Saturazione. Inoltre, ad esempio, è interessante notare la posizione degli Emirati Arabi Uniti (ARE), che possiamo individuare nel punto verde isolato più in alto di tutti, indica che tale paese stia crescendo molto più degli altri in considerazione del livello delle importazioni normalizzate ormai raggiunto.

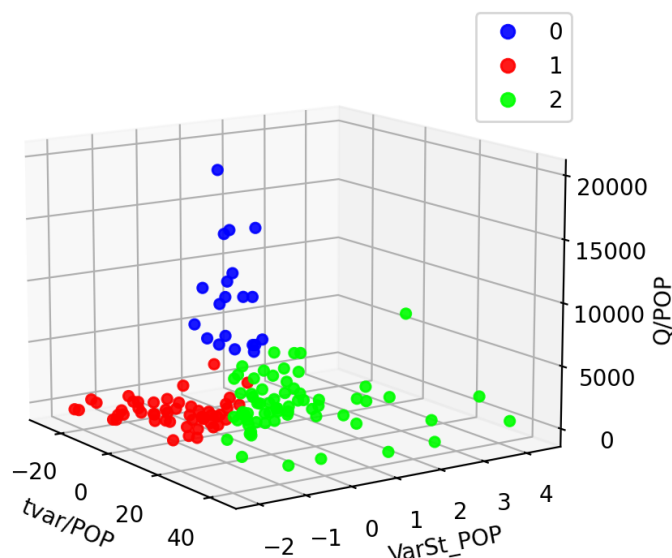


Figura 3.3: Risultati tridimensionali della Cluster Analysis. (0=Saturazione, 1=Latenza, 2=Crescita)

3.1.1 Risultati Analisi

In questa sezione si presenteranno singolarmente i vari Cluster, così da poterne apprezzare le caratteristiche principali da un punto di vista grafico.

Paesi in Latenza

Il primo gruppo che analizzeremo è quello contenente i paesi, le cui importazioni di Lavastoviglie appartengono allo stadio di Latenza. In Figura 3.4 possiamo osservare l'andamento delle serie storiche rappresentanti le *importazioni normalizzate* di tutti i paesi del Cluster, ad eccezione del Regno Unito (GBR). Che, come vedremo dettagliatamente nella sezione 3.1.2, presenta un valore delle *importazioni normalizzate* troppo elevato per essere compatibile con il gruppo dei paesi in Latenza.

Dall'osservazione del grafico possiamo notare come, ad esclusione di qualche picco riscontrabile nella serie storica di Qatar, Argentina e Iran, la gran parte dei paesi mostra un andamento stazionario ad un livello contenuto

($Q/POP < 500$), non mostrando alcuna tendenza all'aumento negli anni più recenti del nostro campione.

Tali livelli, riscontrabili nella maggioranza dei paesi del cluster, ci permettono di considerare con un buon margine di sicurezza tale classificazione come soddisfacente, poiché conferma le aspettative che avevamo, in merito alle caratteristiche delle variabili appartenenti a questo gruppo.

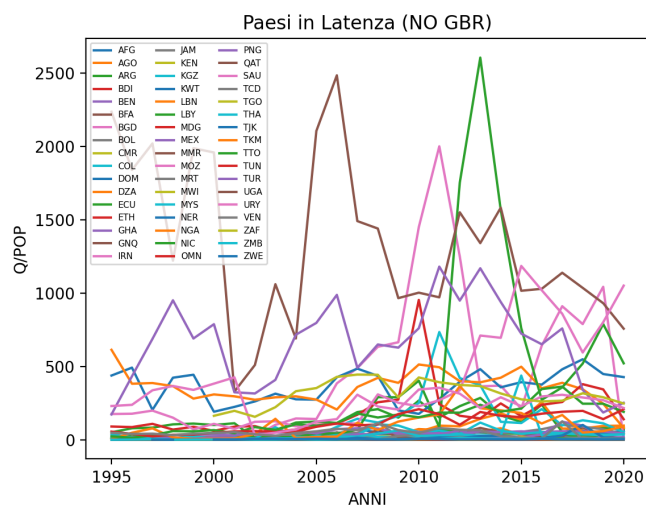


Figura 3.4: Cluster paesi in Latenza senza Regno Unito

Paesi in Crescita

Situazione ben diversa si osserva in Figura 3.5⁹. Tale grafico è tuttavia un grafico parziale, poiché, pur essendo la gran parte dei paesi che sono stati rimossi effettivamente in uno stadio di Crescita, il loro trend non sarebbe stato ancora apprezzabile nella figura. Infatti, si tratta di paesi che stanno iniziando adesso a percorrere lo stadio di Crescita e che non si sono ancora discostati troppo da quel livello "base" che abbiamo visto anche nei paesi in Latenza.

Le differenze con la precedente figura (3.4) appaiono evidenti, sono infatti riscontrabili importanti cambi di inclinazione, che in precedenza non si registravano né a livello grafico né a livello analitico. Nella Figura 3.5 possiamo vedere come quasi tutti i paesi hanno un trend crescente, ed in particolare negli anni seguenti al 2015. Dandoci così la conferma grafica di una corretta classificazione.

Una menzione particolare merita il caso degli Emirati Arabi Uniti (ARE), che nella figura 3.5 si vede svettare sopra tutti con una crescita molto sostenuta.

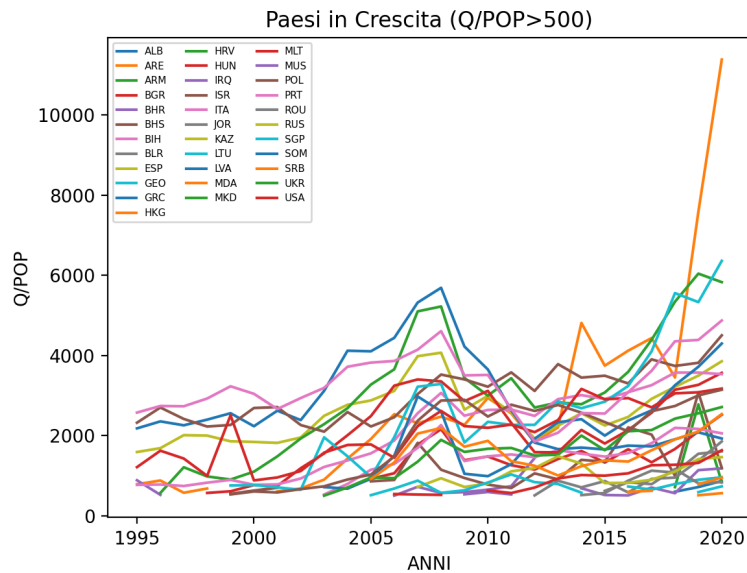


Figura 3.5: Cluster Paesi in Crescita con $Q/POP > 500$

⁹In Appendice (K) può essere osservato il grafico dell'andamento delle importazioni normalizzate di tutti i paesi classificati come "Crescita"

Oltre agli Emirati Arabi esistono altri paesi che, considerando l'obiettivo di questo lavoro, meritano una maggiore considerazione. Dall'analisi specifica del gruppo dei paesi in Crescita, possiamo individuarne un sottogruppo, il cui andamento può essere osservato in Figura 3.6, i cui paesi sono caratterizzati da una crescita maggiore rispetto agli altri, candidandosi quindi come i paesi più vantaggiosi per un eventuale ingresso da parte di un esportatore estero.

In Tabella 3.2 vediamo i valori puntuali delle variabili dei paesi del sottogruppo dei paesi in Crescita, e immediatamente notiamo gli alti valori del tasso di variazione ($tvar/POP$) e della variazione storica ($VarSt_POP$). Ovviamente tutti ben al di sopra della media del Cluster. Da notare il caso dell'Armenia (ARM), che ha raggiunto un picco molto elevato nel 2019 per poi precipitare nel 2020, crollo probabilmente influenzato dalla combinazione tra la guerra Nagorno Karabakh e dalla presenza del Covid19 che hanno raggiunto il loro culmine nel medesimo momento, il terzo trimestre del 2020. Altra considerazione riguarda il volume delle importazioni della Serbia (SRB) e della Bulgaria (BGR), che avevano iniziato il loro percorso di crescita già nel 2004, prima di subire il crollo causato dalla crisi del 2008 e dalla crisi dei debiti sovrani del 2012. Poi, dal 2014, entrambi i paesi sono ritornati sul percorso di crescita che ancora oggi stanno seguendo.

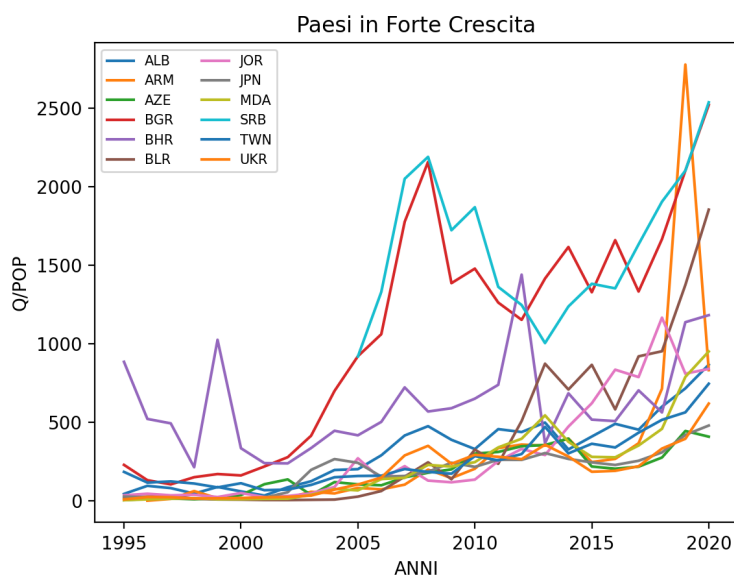


Figura 3.6: Paesi che mostrano una Crescita molto sostenuta

Tabella 3.2: Valori variabili paesi Fig. 3.6

Paesi	Q/POP	tvar/POP	VarSt_POP
ALB	726.74	12.32	1.39
ARM	1440.60	30.16	4.19
AZE	375.50	3.58	2.07
BLG	2095.78	7.47	1.26
BHS	960.10	12.56	2.42
BLR	1393.00	10.80	1.29
JOR	937.89	16.55	-0.13
JPN	401.74	8.07	1.79
MDA	732.01	13.47	3.13
SRB	2180.87	12.59	1.09
TWN	607.44	10.21	0.96
UKR	448.29	10.99	2.74

Paesi in Saturazione

Forse la figura relativa a questo gruppo di paesi è la più rappresentativa. Infatti, osservando Figura 3.7 possiamo vedere come la crescita di quasi tutti questi paesi si sia interrotta intorno al 2009, per poi proseguire in un percorso tendenzialmente stazionario fino al 2020. In pratica, questo andamento suggerisce come l'impatto delle crisi del 2008 e del 2012 sia stato così rilevante da causare, forzatamente, tale livello di saturazione. È come se, per i paesi che crescevano nei primi anni 2000, il livello massimo raggiunto prima delle crisi sia diventato anche il loro livello massimo di Saturazione, sarà quindi interessante vedere nei prossimi anni cosa succederà a quei paesi che stanno crescendo oggi.

Interessante notare come non esista un effettivo valore delle *importazioni normalizzate* (Q/POP) sopra il quale si possa considerare un paese nello stadio di Saturazione. È invece la combinazione delle variabili ad identificare tale stadio. È evidente notare infatti, come il livello di importazioni in cui ogni paese si è stabilizzato risulta essere molto diverso tra i vari stati, indicando l'inesistenza di una soglia univoca in grado di individuare tale stadio. Ad esempio, la Svezia (SWE) ha raggiunto il suo massimo a circa 20.000 \$/mln ab. mentre la Francia (FRA) si è stabilizzata intorno a 6000. Grazie a questo grafico si può effettivamente apprezzare quanto fondamentale e imprescindibile sia l'utilizzo combinato delle tre variabili. Questo perchè, con il solo utilizzo delle importazioni normalizzate potremmo essere indotti a considerare in Crescita paesi con livelli di *importazioni normalizzate* molto distanti dai paesi con i valori massimi.

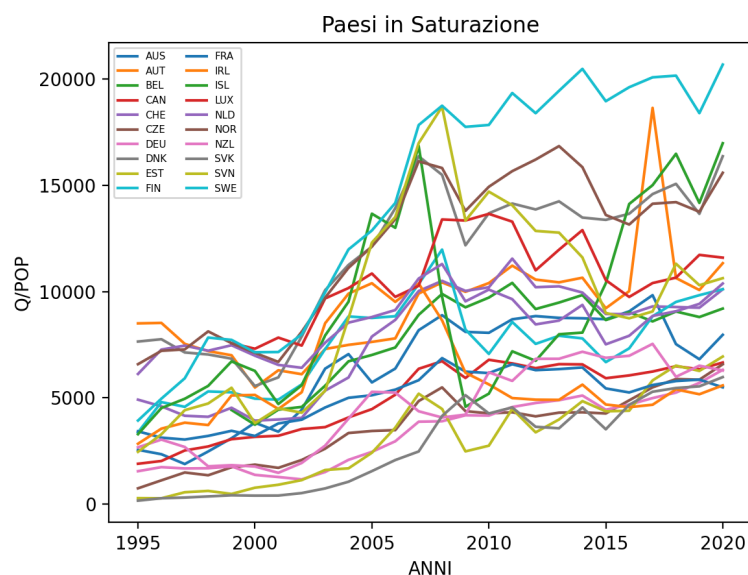


Figura 3.7: Cluster paesi in Saturazione

3.1.2 Criticità analisi: gli errori di classificazione

Ovviamente nel corso dell'esecuzione non tutti i paesi sono stati classificati correttamente, ossia in linea con le aspettative che avevamo dalla teoria, confermate dai valori medi delle variabili di ogni Cluster. Questi errori di classificazione potrebbero essere imputabili a diverse cause, quali valori estremi o viziati nelle variabili, causati da shock endogeni o esogeni al paese.

Regno Unito

Uno degli errori di classificazione più rilevante è stato commesso nell'inclusione del Regno Unito (GBR) nel gruppo dei paesi in Latenza. Il Regno Unito, infatti, presentando valori negativi sia nel *tasso di variazione CAGR* (-0.7538) sia nella *Variazione Storica* (-0.088). Per questo avrebbe più similarità con i paesi in Latenza rispetto a quelli in Saturazione, dove anche se la *Variazione Storica* è per molti paesi negativa il tasso CAGR è solitamente maggiore di zero. Tale particolarità può essere in parte attribuita agli effetti della Brexit, che sommandosi a ciò che si è visto anche in molti altri paesi europei dove le importazioni sono ormai stazionarie dal 2012, nel Regno Unito hanno iniziato negli ultimi anni anche a decrescere. In Figura 3.8 si osserva come il Regno Unito nel complesso abbia caratteristiche ben diverse rispetto agli altri paesi in Latenza.

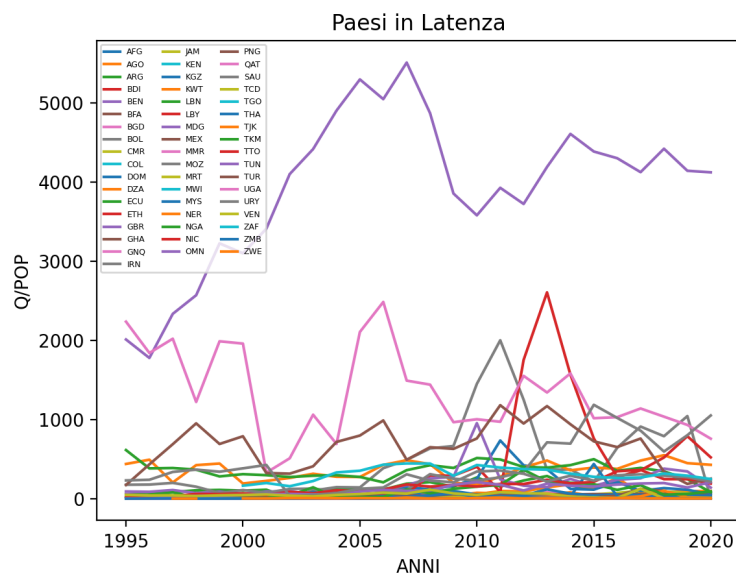


Figura 3.8: Differenza del Regno Unito rispetto agli altri paesi classificati come in Latenza

Germania, Estonia e Repubblica Ceca

Anche altri paesi sono stati classificati erroneamente. In questo caso la principale causa che ha portato ad un errata classificazione è da attribuire al valore della *Variazione Storica*. Come si può osservare anche dalla Figura 3.9, la crescita per questi paesi è iniziata nei primi anni 2000, e questo comporta che, data la lunga durata di questa fase, nel calcolo della *Variazione Storica* non si rilevi quel cambio di inclinazione che invece è tipico dei paesi che entrano in questo stadio. Per cui, per questi tre paesi il valore assunto da tale variabile è relativamente basso (inferiore a 0.6 per tutte e tre), questo perchè essendo calcolata come rapporto tra il volume delle importazioni negli ultimi 3 anni rispetto ai precedenti 10 ed essendo la crescita iniziata più di 13 anni fa, il numeratore risulta decisamente più contenuto rispetto al denominatore. In conclusione, questo ha portato a classificarli come stati in Saturazione, quando in realtà sono ancora in Crescita.

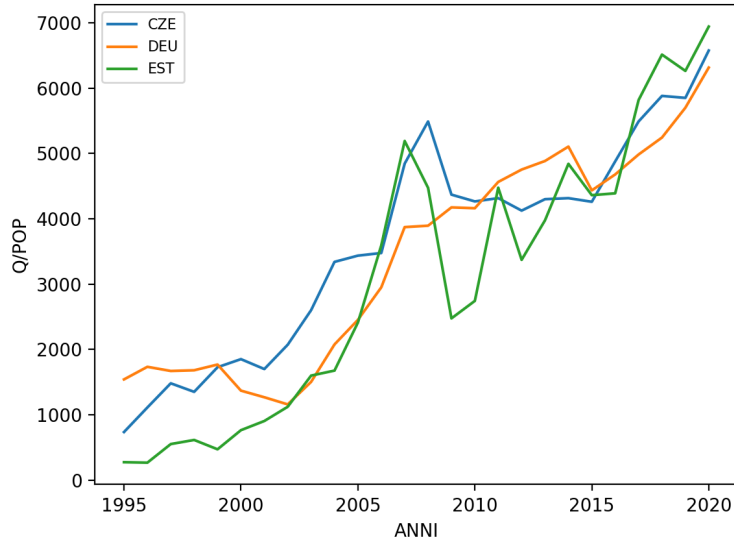


Figura 3.9: Germania, Estonia e Repubblica Ceca

3.2 Linear Discriminant Analysis

Passiamo adesso ad analizzare i risultati ottenuti dal migliore degli algoritmi di Machine Learning supervisionato che abbiamo testato¹⁰. Come vedremo nel corso di questo paragrafo, pur essendo stato l'algoritmo di ML supervisionato più efficiente, l'analisi non è stata così soddisfacente, si sono registrate alcune criticità importanti che abbiamo comunque cercato di comprendere e argomentare qui di seguito. In Figura 3.10 possiamo vedere i risultati della Linear Discriminant Analysis. Appare evidente fin dal primo sguardo l'enorme differenza con i risultati della Cluster. Si può vedere infatti, come, mediante questa tecnica si sia stati in grado di differenziare solamente fra paesi in Latenza e paesi in Crescita, mentre è completamente fallito il riconoscimento dei paesi in Saturazione, che, come si può vedere in figura, risultano sparsi ovunque.

3.2.1 Risultati Analisi

In Tabella 3.3 possiamo osservare i valori medi delle variabili nei vari gruppi selezionati dalla LDA. I paesi in Crescita appaiono ben rappresentati come

¹⁰Se confrontiamo i risultati con la Cluster Analysis, la LDA ha classificato circa il 70% dei paesi correttamente, gli altri algoritmi "supervisionati" hanno avuto rendimenti inferiori al 55%

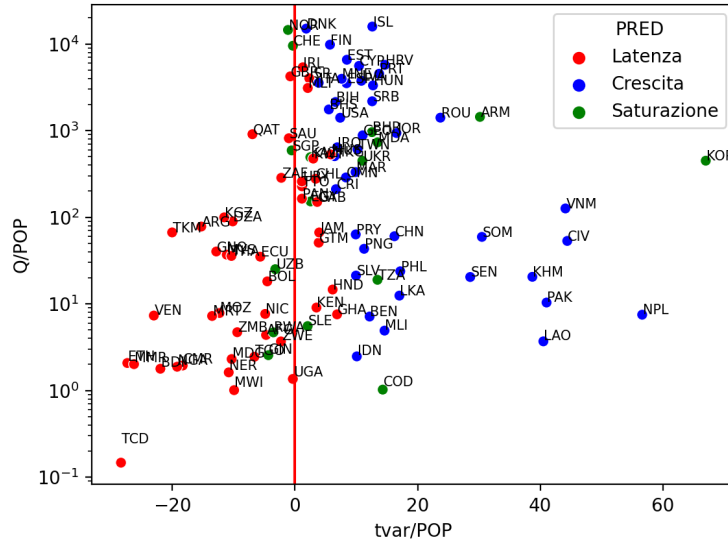


Figura 3.10: Classificazione LDA in scala log

numero, e i valori medi delle variabili sono coerenti con quanto ci saremmo aspettati. Da sottolineare il basso valore della *Variazione Storica*, probabilmente influenzato al ribasso dalla presenza di paesi in realtà in Saturazione ma erroneamente classificati come Crescita, sulle cui cause approfondiremo nel paragrafo seguente. Ricordiamo che alcuni dei paesi più rappresentativi dei vari stadi, essendo stati utilizzati in fase di apprendimento¹¹, e quindi dotati manualmente dell'etichetta relativa al gruppo, come ad esempio la Svezia o gli Emirati Arabi, non sono considerati nella Tabella 3.3. L'inclusione di questi nella tabella avrebbe distorto i valori medi della classificazione LDA, compromettendone la validità e la capacità di individuare eventuali errori. Da un punto di vista di valori medi, anche i paesi in Latenza sembrerebbero correttamente classificati, i valori di tutte le variabili risultano essere i più bassi rispetto a quelli degli altri gruppi, indicando così una certa coerenza con le aspettative iniziali.

Discorso diverso per i paesi in Saturazione. In questo caso la classificazione è stata totalmente fallace e i valori medi di tutte le variabili risultano incoerenti sia con le aspettative, sia con le risultanze della Cluster Analysis. Osservando la Tabella in appendice (L), si può vedere come non solo, i paesi classificati come in Saturazione in realtà non lo siano affatto, ma che i paesi che si troverebbero in uno stadio di Saturazione, siano stati invece classifica-

¹¹In Appendice G possono essere consultati tutti i paesi forniti in apprendimento

ti come in Crescita o Latenza in base all'andamento del *tasso di Variazione CAGR* degli ultimi anni.

Tabella 3.3: Valori medi delle variabili in base al gruppo di appartenenza

	N.Paesi	Q/POP	tvar/POP	VarSt_POP
Latenza	46	468,22	-6,79	-0,43
Crescita	42	2265,49	16,33	0,48
Saturazione	16	1836,88	9,75	3,53

3.2.2 Criticità Analisi: gli errori di classificazione

Come abbiamo affermato all'inizio di questo paragrafo, la principale criticità di quest'ultimo algoritmo è stato il fallimento nell'identificare i paesi in Saturazione.

Osservando Figura 3.11 possiamo farci un'idea di quella che potrebbe essere la ragione principale di questo fallimento. Nel grafico possiamo vedere la distribuzione dei paesi secondo il *tasso di variazione (tvar/POP)* e la *variazione storica (VarSt_POP)*, da cui emerge che la separazione dei gruppi sia molto più netta rispetto alla Fig. 3.10. Questo può indicare che l'algoritmo di LDA abbia riscontrato un'efficienza di classificazione migliore attribuendo un peso maggiore alle variabili appena citate, andando così a sottostimare l'effetto delle *importazioni normalizzate*.

Tale problematica può risiedere direttamente nelle caratteristiche distributive delle *importazioni normalizzate*, che presenta un'ampiezza totale decisamente maggiore rispetto alle altre, ed in particolare, quella del gruppo dei paesi in Saturazione è estraneamente elevata (da 5000 \$/mln ab. a 20000 \$/mln. ab). Tale situazione potrebbe aver indotto l'algoritmo ad attribuire un maggiore rilevanza alle altre due variabili, andando così a perdere la capacità di individuare le principali caratteristiche dei paesi in Saturazione.

Saturazione

Cerchiamo ora di osservare cosa succede all'interno dei vari gruppi. Se ancora c'erano dubbi sulla corretta classificazione o meno dei paesi in Saturazione, possiamo dire che la Figura 3.12 può definitivamente eliminarli tutti. Risulta evidente come, tra i 16 paesi così classificati, solamente due (Norvegia e Svizzera) appartengono davvero a questa classe, mostrando infatti gli stessi andamenti che si erano visti nello stesso gruppo della Cluster Analysis, ossia

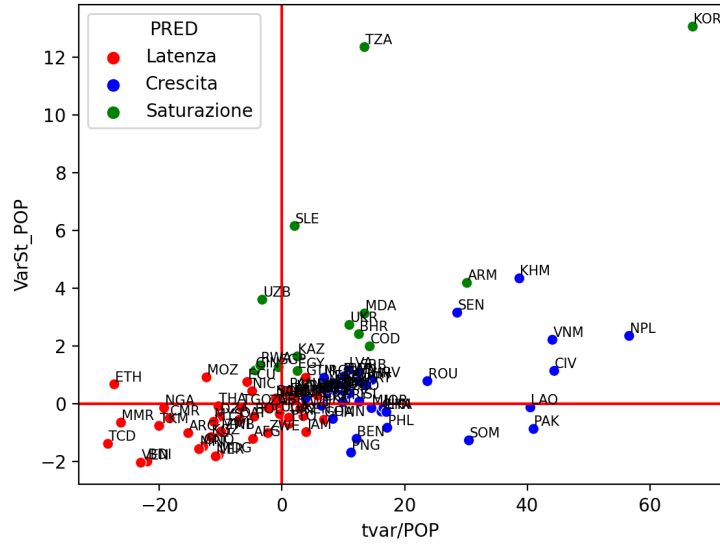


Figura 3.11: Classificazione LDA

una crescita fino al 2009 e un andamento piuttosto stazionario a seguire. Gran parte degli altri paesi sarebbe dovuta essere classificati come in Latenza, ad eccezione di un limitato sottogruppo di paesi formato da Armenia, Singapore, Bahrein e Moldavia, i quali mostrando una leggera tendenza alla crescita e i più alti valori di Q/POP del gruppo (esclusi Norvegia e Svizzera) avrebbero avuto le caratteristiche per essere classificati come paesi in Crescita.

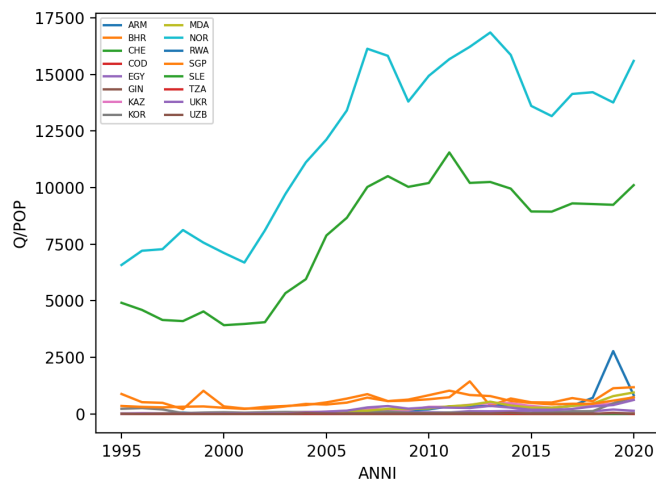


Figura 3.12: Paesi Classificati in Saturazione da LDA

Latenza e Crescita

Per quanto riguarda i paesi in Crescita e Latenza, la situazione appare molto diversa rispetto a quanto osservato finora riguardo ai paesi in Saturazione. Osservando Figura 3.13, possiamo vedere come, per la gran parte dei paesi, la tendenza negli ultimi anni sia effettivamente al rialzo andando così a confermare lo stato di Crescita. Tuttavia, risulta evidente la presenza di paesi nella parte alta del grafico appartenenti, secondo la Cluster Analysis, allo stadio di Saturazione (Danimarca, Islanda e Finlandia).

Chiudiamo con i paesi in Latenza osservabili in Figura 3.14. Anche in questo caso vale il discorso fatto all'inizio di questa sezione e di quanto detto per i paesi in Crescita. Gran parte degli stati è stata infatti classificata correttamente, possiamo vedere come meno di 10 stati disponga di valori $Q/POP > 600$, dove però, solamente 4 di questi, non sono effettivamente riconducibili a questo stadio, ma più probabilmente, visto l'andamento, ad uno stadio di Saturazione (Irlanda, Regno Unito, Malta, Israele).

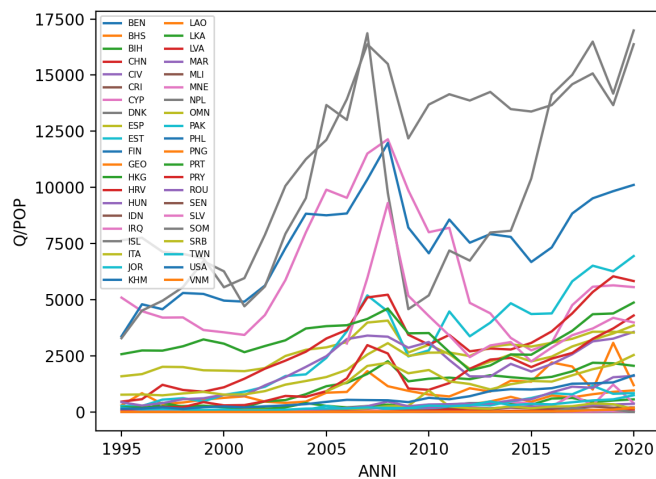


Figura 3.13: Paesi Classificati in Crescita da LDA

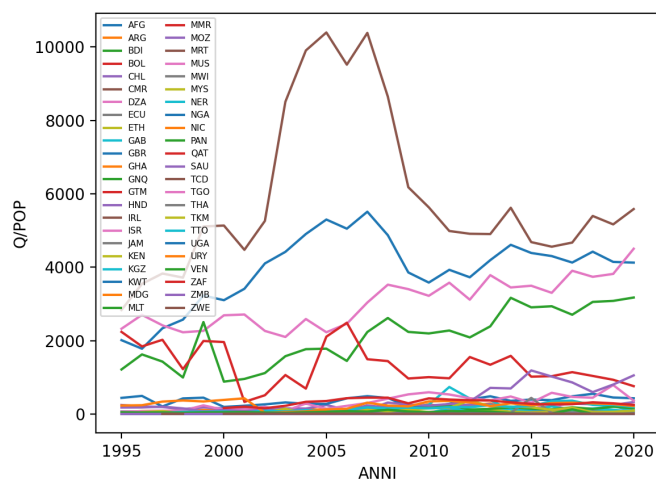


Figura 3.14: Paesi Classificati in Latenza da LDA

3.3 Implicazioni per l'esportatore italiano

In considerazione dei risultati sopra descritti, riteniamo fondamentale dare alcune indicazioni sulle implicazioni di questo lavoro, così da riuscire a rispondere a ciò che è il fulcro del progetto, ossia l'individuazione di paesi in cui esportare il nostro prodotto può rivelarsi una scelta vantaggiosa.

In Tabella 3.4 possiamo osservare una serie di paesi che, data la loro importanza da un punto di vista economico, potrebbero trarre in inganno e portare un potenziale investitore a decidere di esportare i propri prodotti in queste realtà. Tuttavia, grazie alla combinazione delle variabili e all'esecuzione della Cluster Analysis, siamo in grado di dire come tali paesi non siano vantaggiosi, essendo il mercato già saturo. Menzione speciale per la Cina che, pur avendo le importazioni in uno stadio di Crescita¹², ha una forte produzione interna, cosa che rende difficile l'ingresso dall'esterno, infatti pur essendo in presenza di un *tasso di variazione* molto elevato, la *variazione storica* è negativa, indicando come pur crescendo, negli ultimi anni non hanno avuto quell'accelerazione che ci saremmo aspettati da un paese in uno stadio di Crescita.

Tabella 3.4: Paesi in cui non è consigliato esportare

Paese	Classificazione	Q/POP	tvar/POP	VarSt_POP
Svezia	Saturazione	19757.2	0.1195	-0.0375
Lussemburgo	Saturazione	11328.7	-0.7500	0.3380
Francia	Saturazione	5711.7	-1.1639	0.0498
Belgio	Saturazione	9021.8	-0.6285	0.0352
Danimarca	Saturazione	15038.1	1.8384	0.1228
Norvegia	Saturazione	14525.1	-1.1282	0.1687
Cina	Crescita/Latenza	60.4	16.2257	-0.2614

In tabella 3.5, possiamo osservare quei paesi dove invece esportare è assolutamente consigliato. Alla luce anche delle considerazioni fatte nei paragrafi precedenti, potremmo identificare nella zona geografica dell'Est Europa uno dei centri attuali più profittevoli dove importare Lavastoviglie. La gran parte dei paesi in questa zona presenta infatti, tassi di crescita molto elevati che non accennano a diminuire, con valori di *importazioni normalizzate* che, se confrontati con i paesi in Saturazione, indicano ampi spazi per l'esportatore in cui poter crescere ancora.

¹²Combinazione nel valore delle variabili molto al limite, tant'è che in diverse situazioni nell'esecuzione dell'algoritmo di Cluster Analysis il paese veniva incluso a volte nel gruppo dei paesi in Latenza altre volte in quello di Crescita

Tabella 3.5: Paesi la cui situazione è favorevole ad avviare un processo di esportazione

Paese	Classificazione	Q/POP	tvar/POP	VarSt_POP
Albania	Crescita	726.7	12.3230	1.3856
Armenia	Crescita	1440.6	30.1641	4.1891
Bielorussia	Crescita	1393.0	10.7966	1.2829
Bulgaria	Crescita	2095.8	7.4683	1.2559
Emirati Arabi Uniti	Crescita	7472.4	14.6001	3.4356
Lituania	Crescita	5478.62	15.4795	0.9296
Moldavia	Crescita	732.0	13.4686	3.1316
Serbia	Crescita	2180.9	12.5925	1.0916

Conclusioni

Lo scopo di questo lavoro era quello di verificare se, tramite l'utilizzo di strumenti di Machine Learning, fosse possibile individuare gli stadi di sviluppo delle importazioni di un prodotto in un determinato paese.

In considerazione dei risultati ottenuti e ampiamente descritti nelle parti precedenti, possiamo considerare l'algoritmo di Cluster Analysis come l'opzione migliore nel risolvere questo problema. Si è potuto vedere infatti come tale procedura sia in grado di allocare in maniera piuttosto netta e accurata i vari paesi nei tre stadi, riuscendo ad individuare gruppi le cui caratteristiche risultano coerenti con le aspettative che sorgono dall'analisi teorica di base.

In virtù di quanto osservato, possiamo procedere con la seguente constatazione. Se è vero che i vari gruppi identificati dall'algoritmo rispecchiano le caratteristiche attese, è anche vero che tali aspettative sono frutto di interpretazioni di teorie diverse, ricordiamo infatti che, non esistendo una teoria del ciclo vitale delle importazioni, abbiamo dovuto analizzare la teoria del ciclo vitale del prodotto (PLC) e la teoria del ciclo delle esportazioni, cercando di tradurre le informazioni necessarie per strutturare il ciclo vitale delle importazioni. Una volta identificata la struttura, basandoci sempre sulle altre teorie del ciclo vitale a disposizione, abbiamo cercato di identificare quali sarebbero potute essere le caratteristiche di ognuna delle fasi. Quindi, il fatto che la successiva Cluster Analysis abbia dato dei risultati coerenti con le aspettative che avevamo, ci fornisce una doppia conferma, una sui risultati stessi e l'altra sulle ipotesi che avevamo costruito.

Potremmo quindi concludere che, le ipotizzate tre fasi del ciclo vitale delle importazioni (Latenza, Crescita e Saturazione) trovano un riscontro tangibile nell'analisi svolta, dove l'analisi della Cluster Analysis può portare all'implementazione di strategie di marketing mix più accurate e precise, in particolare per quello che concerne la scelta del paese in cui iniziare ad esportare il nostro prodotto.

Bibliografia

Baran P. (1961), *The stages of economic growth*

Hitt M., Duane R. e Hoikisson R, *Strategic Managment: Competitiveness and Globalization (IV Edition)*

Hollensen S., *Global Marketing*

Kassambara A. (2017), *Pratical Guide to principal component methods*

Lancaster G. and Wesenlund I. (1984), *A product life cycle theory for international trade: an empirical investigation*

Meier G. (1989), *Sequence of stages* Prasad Y. (2020), *Business life cycle: 6 stages and their importance*

Raschka S. and Mirjalill V. (2020), *Machine learning con Python*

Rostow W.W. (1959), *The Economic history review*

Rostow W.W. (1953), *The process of economic growth*

Rostow W.W. (1956), *The take-off into self-sustained growth*

Seev H. (1967), *Location of Industry and International Competitiveness* Taddy M., *Business Data Science: Combining Machine Learning and Economics to Optimize, Automate and Accelerate Business Decisions*

Vernon R. (1966), *International Investments and International Trade in the product growth*

Wang Z. (2006), *Learning, Diffusion, and Industry Life Cycle*

Wells L. T.(1968),*A product life cycle for international trade*

Wohlner R. (2020), *Industry Life Cycle*

Yuen Poh L. (1970), *Product life cycle and export competitiveness of the UK electronics industry*

Appendici

Appendice A

Lista paesi banche dati

Tabella A.1: Paesi presenti nei vari dataset

Prodotto	Popolazione	PIL	Dati Utilizzati
AFG	AFG	AFG	AFG
AGO	AGO	AGO	AGO
ALB	ALB	ALB	ALB
ARE	ARE	ARE	ARE
ARG	ARG	ARG	ARG
ARM	ARM	ARM	ARM
AUS	AUS	AUS	AUS
AUT	AUT	AUT	AUT
AZE	AZE	AZE	AZE
BDI	BDI	BDI	BDI
BEL	BEL	BEL	BEL
BEN	BEN	BEN	BEN
BFA	BFA	BFA	BFA
BGD	BGD	BGD	BGD
BGR	BGR	BGR	BGR
BHR	BHR	BHR	BHR
BHS	BHS	BHS	BHS
BIH	BIH	BIH	BIH
BLR	BLR	BLR	BLR
BOL	BOL	BOL	BOL
BRA	BRA	BRA	BRA
CAN	CAF	CAF	CAN
CHE	CAN	CAN	CHE
CHL	CHE	CHE	CHL

Continua nella pagina successiva

Tabella A.1 – *Continua dalla pagina precedente*

Prodotto	Popolazione	PIL	Dati Utilizzati
CHN	CHL	CHL	CHN
CIV	CHN	CHN	CIV
CMR	CIV	CIV	CMR
COD	CMR	CMR	COD
COL	COD	COD	COL
CRI	COL	COL	CRI
CUB	CRI	CRI	CYP
CYP	CYP	CYP	CZE
CZE	CZE	CZE	DEU
DEU	DEU	DEU	DNK
DNK	DNK	DNK	DOM
DOM	DOM	DOM	DZA
DZA	DZA	DZA	ECU
ECU	ECU	ECU	EGY
EGY	EGY	EGY	ESP
ESP	ERI	ERI	EST
EST	ESP	ESP	ETH
ETH	EST	EST	FIN
FIN	ETH	ETH	FRA
FRA	FIN	FIN	GAB
GAB	FRA	FRA	GBR
GBR	GAB	GAB	GEO
GEO	GBR	GBR	GHA
GHA	GEO	GEO	GIN
GIN	GHA	GHA	GNQ
GNQ	GIN	GIN	GRC
GRC	GNQ	GNQ	GTM
GTM	GRC	GRC	HKG
HKG	GTM	GTM	HND
HND	HKG	HKG	HRV
HRV	HND	HND	HUN
HTI	HRV	HRV	IDN
HUN	HTI	HTI	IND
IDN	HUN	HUN	IRL
IND	IDN	IDN	IRN
IRL	IND	IND	IRQ
IRN	IRL	IRL	ISL
IRQ	IRN	IRN	ISR

Continua nella pagina successiva

Tabella A.1 – *Continua dalla pagina precedente*

Prodotto	Popolazione	PIL	Dati Utilizzati
ISL	IRQ	IRQ	ITA
ISR	ISL	ISL	JAM
ITA	ISR	ISR	JOR
JAM	ITA	ITA	JPN
JOR	JAM	JAM	KAZ
JPN	JOR	JOR	KEN
KAZ	JPN	JPN	KGZ
KEN	KAZ	KAZ	KHM
KGZ	KEN	KEN	KWT
KHM	KGZ	KGZ	LAO
KOR	KHM	KHM	LBN
KWT	KOR	KOR	LBY
LAO	KWT	KWT	LKA
LBN	LAO	LAO	LTU
LBR	LBN	LBN	LUX
LBY	LBR	LBR	LVA
LKA	LBY	LBY	MAR
LTU	LKA	LKA	MDA
LUX	LTU	LTU	MDG
LVA	LUX	LUX	MEX
MAR	LVA	LVA	MKD
MDA	MAR	MAR	MLI
MDG	MDA	MDA	MLT
MEX	MDG	MDG	MMR
MKD	MEX	MEX	MNE
MLI	MKD	MKD	MOZ
MLT	MLI	MLI	MRT
MMR	MLT	MLT	MUS
MNE	MMR	MMR	MWI
MOZ	MNE	MNE	MYS
MRT	MOZ	MOZ	NER
MUS	MRT	MRT	NGA
MWI	MUS	MUS	NIC
MYS	MWI	MWI	NLD
NER	MYS	MYS	NOR
NGA	NER	NER	NZL
NIC	NGA	NGA	OMN
NLD	NIC	NIC	PAK

Continua nella pagina successiva

Tabella A.1 – *Continua dalla pagina precedente*

Prodotto	Popolazione	PIL	Dati Utilizzati
NOR	NLD	NLD	PAN
NPL	NOR	NOR	PER
NZL	NPL	NPL	PHL
OMN	NZL	NZL	PNG
PAK	OMN	OMN	POL
PAN	PAK	PAK	PRT
PER	PAN	PAN	PRY
PHL	PER	PER	QAT
PNG	PHL	PHL	ROU
POL	PNG	PNG	RUS
PRT	POL	POL	RWA
PRY	PRT	PRT	SAU
QAT	PRY	PRY	SEN
ROU	QAT	QAT	SGP
ROW	ROU	ROU	SLV
RUS	RUS	RUS	SOM
RWA	RWA	RWA	SRB
SAU	SAU	SAU	SVK
SDN	SDN	SDN	SVN
SEN	SEN	SEN	SWE
SGP	SGP	SGP	TCD
SLE	SLE	SLE	TGO
SLV	SLV	SLV	THA
SOM	SOM	SOM	TJK
SRB	SRB	SRB	TKM
SSD	SSD	SSD	TTO
SVK	SVK	SVK	TUN
SVN	SVN	SVN	TUR
SWE	SWE	SWE	TWN
SYR	SYR	SYR	UGA
TCD	TCD	TCD	UKR
TGO	TGO	TGO	URY
THA	THA	THA	USA
TJK	TJK	TJK	UZB
TKM	TKM	TKM	VEN
TTO	TTO	TTO	VNM
TUN	TUN	TUN	ZAF
TUR	TUR	TUR	ZMB

Continua nella pagina successiva

Tabella A.1 – *Continua dalla pagina precedente*

Prodotto	Popolazione	PIL	Dati Utilizzati
TWN	TWN	TWN	ZWE
TZA	TZA	TZA	
UGA	UGA	UGA	
UKR	UKR	UKR	
URY	URY	URY	
USA	USA	USA	
UZB	UZB	UZB	
VEN	VEN	VEN	
VNM	VNM	VNM	
YEM	YEM	YEM	
ZAF	ZAF	ZAF	
ZMB	ZMB	ZMB	
ZWE	ZWE	ZWE	

Appendice B

Comandi Python variabile Importazioni normalizzate

```
def CalcoloImportNorm(DFIMPORT, DF_NORM):
    bb=DFIMPORT.reset_index()
    bb=bb.rename(columns={"DIC": "COUNTRY"})
    DF_NORM=DF_NORM.reset_index()
    Defprova=pandas.merge(bb,DF_NORM,on=["COUNTRY", "YEAR"])
    Defprova=Defprova.rename(columns={"IMFWEO": "NORM"})
    Defprova=Defprova[["YEAR", "COUNTRY", "Q", "NORM"]]
    Defprova["Q/NORM"]=Defprova["Q"]/Defprova["NORM"]
    prodDefprova=pandas.DataFrame()
    prodDefprova=Defprova[["YEAR", "COUNTRY", "Q/NORM"]]
    prodDefprova["YEAR"]=prodDefprova["YEAR"].map(str)
    prodDefprova=prodDefprova[prodDefprova["YEAR"]<"2021"]
    paesi=pandas.DataFrame()
    Q_NORM=prodDefprova[["YEAR", "COUNTRY", "Q/NORM"]]
    Q_NORM=Q_NORM.groupby(["YEAR", "COUNTRY"]).sum()
    return(Q_NORM)

def ImportMedio(DFIMPORT, DF_NORM):
    bb=DFIMPORT.reset_index()
    bb=bb.rename(columns={"DIC": "COUNTRY"})
    DF_NORM=DF_NORM.reset_index()
    Defprova=pandas.merge(bb,DF_NORM,on=["COUNTRY", "YEAR"])
    Defprova=Defprova.rename(columns={"IMFWEO": "NORM"})
    Defprova=Defprova[["YEAR", "COUNTRY", "Q", "NORM"]]
    Defprova["Q/NORM"]=Defprova["Q"]/Defprova["NORM"]
```



```
prodDefprova=pandas.DataFrame()  
prodDefprova=Defprova[["YEAR","COUNTRY","Q/NORM"]]  
prodDefprova["YEAR"]=prodDefprova["YEAR"].map(str)  
Q_mean=prodDefprova[prodDefprova["YEAR"]>="2018"]  
Q_mean=Q_mean[Q_mean["YEAR"]<"2021"]  
Q_mean=Q_mean.groupby(["COUNTRY"]).mean().reset_index()  
return(Q_mean)
```

Appendice C

Comandi Python calcolo CAGR

```
def tassoVariazionePIL(Q_prodDef):
    paesi=pandas.DataFrame()
    Q_PIL=Q_prodDef[["YEAR","COUNTRY","Q/PIL"]]
    Q_PIL=Q_PIL.groupby(["YEAR","COUNTRY"]).sum()
    lp1=Q_PIL.reset_index()
    lp1=lp1["COUNTRY"].unique().tolist()
    for i in lp1:
        TOT2=Q_PIL.loc(axis=0)[: , i]
        TOT2["tvar/PIL_4"]=((TOT2["Q/PIL"]/TOT2["Q/PIL"].shift(5))**
            (1/5)-1)
        paesi=pandas.concat([paesi ,TOT2])
    paesiDefPIL=pandas.DataFrame()
    paesiDefPIL=paesi[["Q/PIL","tvar/PIL_4"]]
    paesiDefPIL.dropna(subset=["tvar/PIL_4"], inplace=True)
    return(paesiDefPIL)

def tassoVariazionePOP(Q_prodDef):
    paesi=pandas.DataFrame()
    Q_POP=Q_prodDef[["YEAR","COUNTRY","Q/POP"]]
    Q_POP=Q_POP.groupby(["YEAR","COUNTRY"]).sum()
    lp1=Q_POP.reset_index()
    lp1=lp1["COUNTRY"].unique().tolist()
    for i in lp1:
        TOT2=Q_POP.loc(axis=0)[: , i]
        TOT2["tvar/POP_4"]=((TOT2["Q/POP"]/TOT2["Q/POP"].shift(5))**
            (1/5)-1)*100
        paesi=pandas.concat([paesi ,TOT2])
```

```
paesiDefPOP=pandas.DataFrame()  
paesiDefPOP=paesi[['Q/POP', 'tvar/POP_4']].reset_index()  
paesiDefPOP.dropna(subset=["tvar/POP_4"], inplace=True)  
return(paesiDefPOP)
```

Appendice D

Comandi Python Tasso Storico

```
def TassoVariazioneStorica(DATA):
    paesiTasSt=pandas.DataFrame()
    lpaesi=DATA.reset_index()
    lpaesi=lpaesi["COUNTRY"].unique().tolist()
    for i in lpaesi:
        TOT2=DATA.loc(axis=0)[: , i].reset_index()
        Variabili=pandas.DataFrame()
        Variabili=TOT2[TOT2["YEAR"]=="2020"]
        TOT2=TOT2.set_index("YEAR")
        if "2016" in list(TOT2.reset_index()["YEAR"]):
            if "2017" in list(TOT2.reset_index()["YEAR"]):
                if "2018" in list(TOT2.reset_index()["YEAR"]):
                    if "2020" in list(TOT2.reset_index()["YEAR"]):
                        if "2008" in list(TOT2.reset_index()["YEAR"]):
                            if "2007" in list(TOT2.reset_index()["YEAR"]):
                                if "2006" in list(TOT2.reset_index()["YEAR"]):
                                    M3=((TOT2.loc["2016", "Q/NORM"]+TOT2.loc["2017", "Q/NORM"]+
                                        TOT2.loc["2018", "Q/NORM"])/3)
                                    M10=((TOT2.loc["2006", "Q/NORM"]+TOT2.loc["2007", "Q/NORM"]+
                                        TOT2.loc["2008", "Q/NORM"])/3)
                                    M0=((TOT2.loc["2020", "Q/NORM"]*0.55+
                                        TOT2.loc["2019", "Q/NORM"]*0.45))
                                    Variabili["DM3"]=M0-M3
                                    if (M3-M10)>((M3/25)*10):
                                        Variabili["DM10"]=M3-M10
                                    else: Variabili["DM10"]=(M3/25)*10
                                Variabili["tasSt_norm"]=Variabili["DM3"]/Variabili["DM10"]
```

```
    Variabili=Variabili[["YEAR","COUNTRY","tasSt_norm"]]
    paesiTasSt=pandas.concat([paesiTasSt, Variabili])
paesiTasSt=paesiTasSt.reset_index()
paesiTasSt=paesiTasSt[["YEAR","COUNTRY","tasSt_norm"]]
return(paesiTasSt)
```

Appendice E

Dimostrazione applicazione formule Variabile Storica

Come possiamo osservare nella Figure E.1 e E.2, l'utilizzo della Formula (2.3) al solo caso in cui $M3 - M10 < 0$ porta a considerare come paesi in Saturazione (Cluster 2) anche quei paesi (interno al cerchio rosso) che si trovano ancora oggi in uno stadio di forte Crescita(Cluster 0), prima fra tutti gli Emirati Arabi Uniti (ARE).

Nella Figura E.2, dove viene utilizzata la condizione delle formule corretta, possiamo vedere che una parte di quei paesi che precedentemente venivano classificati come in Saturazione adesso risultano, correttamente, nello stato di Crescita.

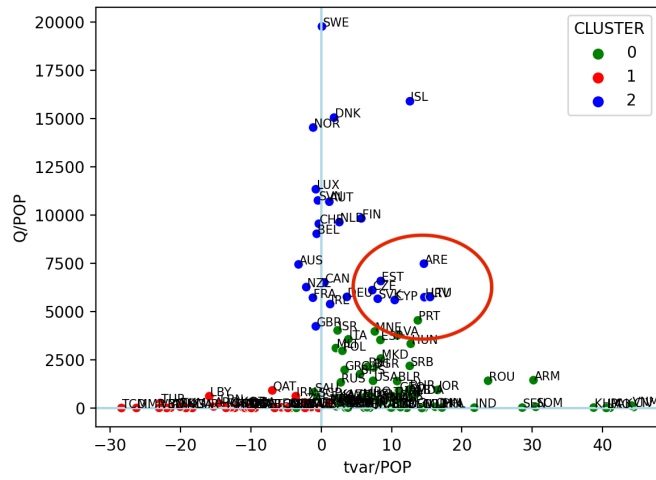


Figura E.1: Cluster Analysis con Variazione Storica calcolata limitando applicazione Formula (2.3) ai soli casi con $M3 - M10 < 0$

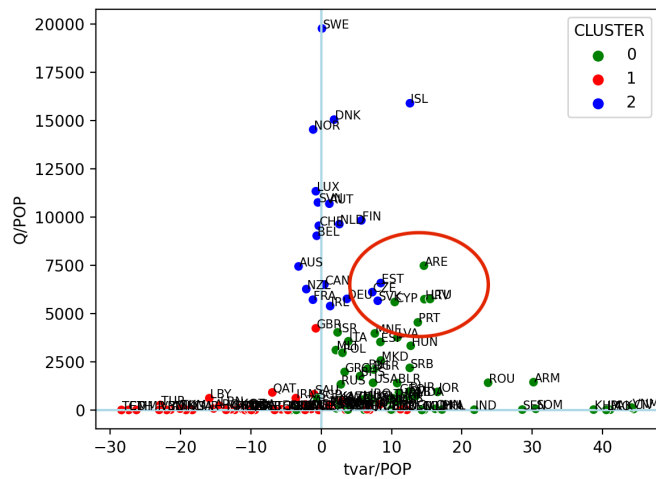


Figura E.2: Cluster Analysis con Variazione Storica calcolata con la procedura corretta

Appendice F

Codice Cluster Analysis

```
def Clustering(ALLFEAT2,NCLUSTER):
    ALLFEATURE=ALLFEAT2.drop('COUNTRY',axis=1)
    ALLFEATURE=StandardScaler().fit_transform(ALLFEATURE)
    FEATURECL=pandas.DataFrame(columns=list(ALLFEAT2['COUNTRY'].
        unique()))
    kmeans=KMeans(n_clusters=NCLUSTER)
    kmeans.fit(ALLFEATURE)
    FEATURECL.loc[0]=kmeans.labels_
    FEATURECL=FEATURECL.stack().to_frame().reset_index()
    FEATURECL=FEATURECL.rename(columns={'level_1':'COUNTRY',
        0:'CLUSTER'}).drop('level_0',axis=1)
    TASSOCL=pandas.merge(FEATURECL, ALLFEAT2, on='COUNTRY')
    return(TASSOCL)
```


Appendice G

Lista Paesi Addestramento ML supervisionato

Tabella G.1: Paesi utilizzati per l'apprendimento

SATURAZIONE	CRESCITA	LATENZA
SWE	ARE	DOM
DNK	CYP	UGA
FRA	CRI	GHA
DEU	SRB	KGZ
EST	IND	TGO
CHE	CHN	KEN
	TWN	ZMB
	ALB	MRT
	UKR	TUR

```
# Random Forest
from sklearn.ensemble import RandomForestClassifier
rdf = RandomForestClassifier().fit(X_train, y_train)
print('Accuracy of Random Forest classifier on training set: {:.2f}'
      .format(rdf.score(X_train, y_train)))
print('Accuracy of Random Forest classifier on test set: {:.2f}'
      .format(rdf.score(X_test, y_test)))
# comando utile: rdf.feature_importances_
```

```

#Accuracy of Decision Tree classifier on training set: 1.00
#Accuracy of Decision Tree classifier on test set: 1.00

#K-Nearest Neighbors
from sklearn.neighbors import KNeighborsClassifier
knn = KNeighborsClassifier(n_neighbors=5)# seleziona
    il numero di distanze da calcolare, default=5
knn.fit(X_train, y_train)
print('Accuracy of K-NN classifier on training set: {:.2f}'.
      format(knn.score(X_train, y_train)))
print('Accuracy of K-NN classifier on test set: {:.2f}'.
      format(knn.score(X_test, y_test)))

#Accuracy of K-NN classifier on training set: 0.93
#Accuracy of K-NN classifier on test set: 1.00

#Linear Discriminant Analysis
from sklearn.discriminant_analysis import LinearDiscriminantAnalysis
lda = LinearDiscriminantAnalysis()
lda.fit(X_train, y_train)
print('Accuracy of LDA classifier on training set: {:.2f}'.
      format(lda.score(X_train, y_train)))
print('Accuracy of LDA classifier on test set: {:.2f}'.
      format(lda.score(X_test, y_test)))

#Accuracy of LDA classifier on training set: 0.93
#Accuracy of LDA classifier on test set: 0.90

#Support Vector Machine
from sklearn.svm import SVC
svm = SVC()
svm.fit(X_train, y_train)
print('Accuracy of SVM classifier on training set: {:.2f}'.
      format(svm.score(X_train, y_train)))
print('Accuracy of SVM classifier on test set: {:.2f}'.
      format(svm.score(X_test, y_test)))

#Accuracy of SVM classifier on training set: 0.97
#Accuracy of SVM classifier on test set: 0.80

```

Appendice H

Cluster Analysis con dati normalizzati su PIL

Tabella H.1: Paesi nello stadio di **Latenza**

Paese	Q/PIL	tvar/PIL	VarSt_PIL
AGO	6,0399	-0,1015	0,4612
ARG	7,7654	-0,0907	-0,4333
AUS	136,3156	-0,0165	-0,3439
BDI	6,6302	-0,2146	-1,9544
BEN	5,7023	0,1119	-1,3475
BFA	1,0677	-0,0548	-1,1588
BGD	0,4684	-0,0068	-1,4702
BHS	54,8974	0,0531	0,6159
BOL	5,2540	-0,0689	-0,7359
CHE	110,1298	-0,0001	0,0874
CHL	19,4742	0,0367	0,5740
CHN	5,9431	0,0948	-0,4331
CMR	1,2812	-0,1902	-0,5987
COL	7,3460	0,0234	0,1221
CRI	17,0621	0,0437	-0,0007
DEU	122,5707	0,0277	0,4362
DOM	6,9647	-0,0026	-0,0620
DZA	23,9413	-0,0503	-0,3118
ECU	5,8411	-0,0532	0,9614
EGY	49,9852	0,0574	0,6268
ETH	2,1956	-0,3319	0,1407
FRA	136,7529	-0,0101	0,0114

Continua nella pagina successiva

Tabella H.1 – *Continua dalla pagina precedente*

Paese	Q/PIL	tvar/PIL	VarSt_PIL
GAB	18,8612	0,0635	0,3147
GBR	100,8177	0,0078	-0,0735
GHA	3,3538	0,0492	-0,7700
GIN	2,5668	-0,1024	0,2224
GRC	104,7761	0,0495	0,3416
GTM	11,8019	0,0031	0,7573
HKG	10,6247	0,0292	-0,1616
HND	5,8578	0,0373	0,1966
IDN	0,6203	0,0727	0,3208
IRL	66,0058	-0,0600	-0,1657
IRN	97,4973	-0,0818	-1,3275
ISR	92,5041	-0,0103	0,1139
ITA	106,8845	0,0420	0,2282
JAM	11,8884	0,0254	-0,9823
JPN	9,9872	0,0684	1,6929
KAZ	52,2341	0,0768	1,4597
KEN	4,4321	-0,0195	-0,8418
KGZ	75,5310	-0,1238	-1,2445
KWT	17,6566	0,0966	0,0683
LBN	34,6736	-0,0932	-0,8506
LBY	131,2433	-0,1738	1,1810
LKA	3,2224	0,1633	-0,1244
MAR	100,6950	0,0878	0,8755
MDG	4,5690	-0,0978	-1,7441
MEX	5,5820	0,0075	-0,2983
MLI	5,4006	0,1262	-0,2343
MLT	100,4071	-0,0162	0,0983
MMR	1,4778	-0,2772	-0,8550
MOZ	15,9492	-0,0726	0,6976
MUS	50,8296	0,0467	0,3751
MYS	3,4245	-0,1163	-0,7022
NER	2,8505	-0,1176	-1,8529
NGA	0,8741	-0,1383	-0,1810
NIC	3,9203	-0,0529	0,8404
NZL	148,8899	-0,0239	-0,1753
OMN	17,5103	0,1215	-0,4719
PAN	11,3220	-0,0143	0,6992
PER	2,7019	0,0310	1,8433

Continua nella pagina successiva

Tabella H.1 – *Continua dalla pagina precedente*

Paese	Q/PIL	tvar/PIL	VarSt_PIL
PHL	7,1248	0,1428	-0,9766
PNG	15,2594	0,1155	-2,0131
PRY	11,9530	0,1175	0,3239
QAT	14,8067	-0,0054	-0,4532
SAU	37,4979	0,0021	0,2185
SGP	9,3324	-0,0284	1,2377
SLV	5,2538	0,0764	0,6050
TGO	2,7177	-0,0812	-0,2790
THA	4,8214	-0,1416	-0,3460
TJK	50,9666	0,1060	-0,9634
TKM	8,8198	-0,1678	-0,9079
TTO	13,8698	0,0481	-0,6208
TUN	54,3717	0,0545	0,0001
TUR	28,5174	-0,1865	-1,3651
TWN	22,6313	0,0652	0,9972
UGA	1,4660	-0,0081	-0,6024
URY	14,8550	0,0144	-0,3580
USA	22,1004	0,0428	0,3624
VEN	3,4255	0,0100	-1,3235
ZMB	3,7479	-0,0463	-0,6045
ZWE	2,5005	-0,0272	-0,8868
AFG	7,4973	-0,0040	-1,6535
RWA	5,7197	-0,0529	1,1019
TCD	0,2111	-0,2370	-1,2431
GNQ	4,3254	-0,0011	-0,8042
LUX	96,9622	-0,0153	0,1942
MWI	1,9016	-0,1205	-1,2423
ZAF	44,5073	-0,0072	0,0236
COD	1,8312	0,1094	1,6169
MRT	3,8376	-0,1384	-1,6422

Tabella H.2: Paesi nello stadio di **Crescita**

Paese	Q/PIL	tvar/PIL	VarSt_PIL
ARE	181,2413	0,1521	3,7256
AZE	82,0200	0,1230	1,8862

Continua nella pagina successiva

Tabella H.2 – *Continua dalla pagina precedente*

Paese	Q/PIL	tvar/PIL	VarSt_PIL
BHR	38,7894	0,1235	2,2559
BRA	11,1474	0,0819	2,7410
CIV	23,8004	0,4118	1,1550
IND	7,5525	0,1563	2,7857
KHM	12,3564	0,2835	5,2602
KOR	14,0875	0,6284	13,0999
NPL	6,3095	0,4664	2,3648
PAK	6,9822	0,3918	-0,8540
SEN	14,1304	0,2622	3,1972
TZA	17,6031	0,1160	13,3632
UZB	14,2964	0,0523	4,5496
VNM	37,0583	0,3577	2,2189
SLE	10,4930	0,0706	7,9254
LAO	1,4306	0,3313	-0,2051

Tabella H.3: Paesi nello stadio di **Saturazione**

Paese	Q/PIL	tvar/PIL	VarSt_PIL
ALB	138,9736	0,0802	0,9530
AUT	213,8153	0,0066	-0,5088
BGR	213,6021	0,0199	0,6749
BIH	351,8128	0,0274	0,0754
BLR	213,5503	0,1339	0,9847
CAN	143,3082	0,0198	0,1082
CYP	198,6082	0,0863	0,5861
CZE	261,3678	0,0320	0,0883
DNK	246,7544	0,0119	0,0135
ESP	121,9073	0,0769	0,7567
EST	283,7788	0,0410	0,1596
FIN	199,7439	0,0503	0,3056
GEO	194,2405	0,1048	0,4081
HRV	393,1563	0,1201	0,6820
HUN	203,8563	0,0854	0,4859
IRQ	116,4146	0,1013	0,9466
ISL	235,6078	0,0701	0,3644
JOR	215,6906	0,1553	-0,1595

Continua nella pagina successiva

Tabella H.3 – *Continua dalla pagina precedente*

Paese	Q/PIL	tvar/PIL	VarSt_PIL
LTU	293,4955	0,1020	0,4836
LVA	212,4465	0,0690	0,8162
MDA	165,2422	0,0561	2,2704
MKD	425,2752	0,0526	0,2660
NLD	182,4295	0,0149	0,2128
NOR	195,5482	0,0263	0,3621
POL	189,4101	0,0002	0,2522
PRT	197,5408	0,1164	0,5400
ROU	110,7838	0,1673	0,4785
RUS	121,7741	0,0575	1,1626
SVK	293,5174	0,0624	0,1080
SVN	414,4521	-0,0297	0,0011
SWE	374,8459	0,0173	0,0175
UKR	126,2232	0,0675	1,2844
ARM	323,0006	0,2607	3,8524
BEL	194,9591	-0,0112	-0,0156
SOM	176,0308	0,2859	-1,3287
SRB	292,9003	0,0895	0,5779
MNE	469,1036	0,0367	0,4411

Appendice I

Cluster Analysis con dati normalizzati su POP

Tabella I.1: Paesi nello stadio di **Latenza**

Paese	Q/POP	tvar/POP	VarSt_POP
AFG	4,3787	-4,7036	-1,2069
AGO	15,6829	-21,3846	-0,7942
ARG	78,1177	-15,2662	-1,0042
BDI	1,7834	-21,9347	-1,9890
BEN	7,1421	12,1766	-1,2024
BFA	0,8642	-3,7079	-1,0215
BGD	0,8394	9,8407	-1,2541
BOL	18,1832	-4,5061	-0,4373
CHN	60,3756	16,2257	-0,2614
CMR	1,9431	-18,2911	-0,4962
COL	45,4862	-1,6275	-0,1047
DOM	56,4635	3,4602	-0,0621
DZA	89,6007	-10,1359	-0,4337
ECU	35,0484	-5,6525	0,7616
ETH	2,0776	-27,3251	0,6818
GBR	4229,6721	-0,7538	-0,0880
GHA	7,5678	6,8703	-0,5429
GNQ	40,2255	-12,8220	-1,4702
IRN	613,0297	-3,5937	-0,9857
JAM	66,7266	3,9599	-0,9713
KEN	9,0663	3,5110	-0,4189
KGZ	99,6691	-11,5292	-1,1436

Continua nella pagina successiva

Tabella I.1 – *Continua dalla pagina precedente*

Paese	Q/POP	tvar/POP	VarSt_POP
KWT	475,3031	2,9895	-0,1696
LBN	223,9319	-14,4516	-1,4278
LBY	610,4809	-15,8812	1,5462
MDG	2,3068	-10,3180	-1,7654
MEX	52,6329	-1,4119	-0,3289
MMR	2,0147	-26,2318	-0,6425
MOZ	7,8462	-12,2938	0,9232
MRT	7,2488	-13,5021	-1,5594
MWI	1,0094	-9,9143	-0,9527
MYS	36,9839	-11,1097	-0,6141
NER	1,6190	-10,7895	-1,8115
NGA	1,8821	-19,2015	-0,1370
NIC	7,6566	-4,8486	0,4404
OMN	287,8800	8,3159	-0,5146
PNG	43,2235	11,2788	-1,6794
QAT	908,6864	-6,9311	-0,5437
SAU	816,3456	-0,9946	0,1897
TCD	0,1469	-28,3560	-1,3777
TGO	2,4528	-6,6017	-0,1180
THA	35,7655	-10,3528	-0,0777
TJK	42,4596	6,5258	-0,8688
TKM	66,8149	-20,0170	-0,7552
TTO	227,7055	1,1498	-0,6853
TUN	182,3906	1,3408	-0,1441
TUR	259,9047	-22,8363	-1,5468
UGA	1,3605	-0,3399	-0,3532
URY	258,8654	1,1321	-0,4715
VEN	7,3305	-23,0015	-2,0295
ZAF	285,2188	-2,2254	-0,1327
ZMB	4,7091	-9,4019	-0,9583
ZWE	3,6864	-2,2760	-1,0042

Tabella I.2: Paesi nello stadio di **Crescita**

Paese	Q/POP	tvar/POP	VarSt _POP
ALB	726,7385	12,3230	1,3856
ARE	7472,3959	14,6001	3,4356
ARM	1440,6020	30,1641	4,1891
AZE	375,5014	3,5867	2,0743
BGR	2095,7840	7,4683	1,2559
BHR	960,1013	12,5551	2,4168
BHS	1755,6747	5,5323	0,4293
BIH	2137,9529	6,6075	0,3378
BLR	1393,0006	10,7966	1,2829
BRA	86,6179	2,1596	1,4550
CHL	278,1119	3,3834	0,2974
CIV	53,2873	44,3964	1,1461
COD	1,0239	14,3139	1,9951
CRI	210,9852	6,6918	0,0125
CYP	5591,9144	10,4550	0,6287
EGY	152,3137	2,5477	1,1432
ESP	3522,1283	8,4212	0,7208
GAB	150,2323	3,6340	0,3632
GEO	881,7460	11,0399	0,3773
GIN	2,5579	-4,3346	1,1619
GRC	1971,9494	3,3454	0,2787
GTM	50,9270	3,8783	0,9072
HKG	506,5151	6,4627	-0,0599
HND	14,6434	6,1500	0,2663
HRV	5737,3012	14,6653	0,8358
HUN	3325,1455	12,7287	0,7783
IDN	2,4763	10,0891	0,4420
IND	14,9066	21,7580	1,8360
IRQ	641,3931	6,8828	0,9100
ISR	4018,6033	2,3376	0,3747
ITA	3559,1943	3,8725	0,1870
JOR	937,8959	16,5501	-0,1336
JPN	401,7404	8,0756	1,7923
KAZ	493,3368	2,5174	1,6479
KHM	20,5286	38,7109	4,3473
LAO	3,6941	40,5041	-0,1167
LKA	12,4573	17,0348	-0,2745

Continua nella pagina successiva

Tabella I.2 – *Continua dalla pagina precedente*

Paese	Q/POP	tvar/POP	VarSt POP
LTU	5748,6206	15,4795	0,9296
LVA	3756,7816	10,8330	1,1647
MAR	331,8565	9,8858	0,9659
MDA	732,0127	13,4686	3,1316
MKD	2563,5234	8,4762	0,4711
MLI	4,9141	14,6103	-0,1403
MLT	3102,4115	2,0937	0,2021
MNE	3965,1188	7,6275	0,5582
MUS	530,6090	5,8387	0,2941
PAK	10,3639	41,0208	-0,8627
PAN	163,8636	1,1383	0,4535
PER	17,6397	3,8191	1,0972
PHL	23,7771	17,1739	-0,8208
POL	2961,7201	3,0272	0,6007
PRT	4537,2252	13,7356	0,6583
PRY	63,3508	9,9163	0,0727
ROU	1410,5229	23,7170	0,7917
RUS	1326,5655	2,7972	1,3264
RWA	4,6930	-3,5125	1,3357
SEN	20,4228	28,5758	3,1593
SGP	591,0173	-0,5313	1,2731
SLV	21,2156	9,9553	0,6644
SOM	59,3497	30,4867	-1,2589
SRB	2180,8608	12,5925	1,0916
TWN	607,4380	10,2112	0,9608
UKR	448,2866	10,9955	2,7354
USA	1407,7027	7,3780	0,4220
UZB	25,1453	-3,2044	3,6083
VNM	126,3544	44,0994	2,2220

Tabella I.3: Paesi nello stadio di **Saturazione**

Paese	Q/POP	tvar/POP	VarSt POP
AUS	7435,4072	-3,2043	-0,3865
AUT	10683,7979	1,1680	-0,4456
BEL	9021,7655	-0,6285	0,0352

Continua nella pagina successiva

Tabella I.3 – *Continua dalla pagina precedente*

Paese	Q/POP	tvar/POP	VarSt_POP
CAN	6496,3995	0,4587	0,1035
CHE	9541,0174	-0,3276	0,1488
CZE	6103,0274	7,2709	0,3847
DEU	5753,2227	3,6616	0,5382
DNK	15038,1346	1,8384	0,1228
EST	6573,7844	8,4691	0,4772
FIN	9819,2744	5,7124	0,4164
FRA	5711,6612	-1,1639	0,0498
IRL	5381,5109	1,2828	0,2672
ISL	15885,6345	12,6166	0,0843
LUX	11328,6933	-0,7500	0,3380
NLD	9626,0880	2,5895	0,3870
NOR	14524,0697	-1,1282	0,1687
NZL	6260,4673	-2,1089	-0,1641
SVK	5656,8041	8,0440	0,3111
SVN	10748,9905	-0,4550	0,2000
SWE	19757,1959	0,1195	-0,0375

Appendice J

Grafici Cluster (POP) in scala logaritmica

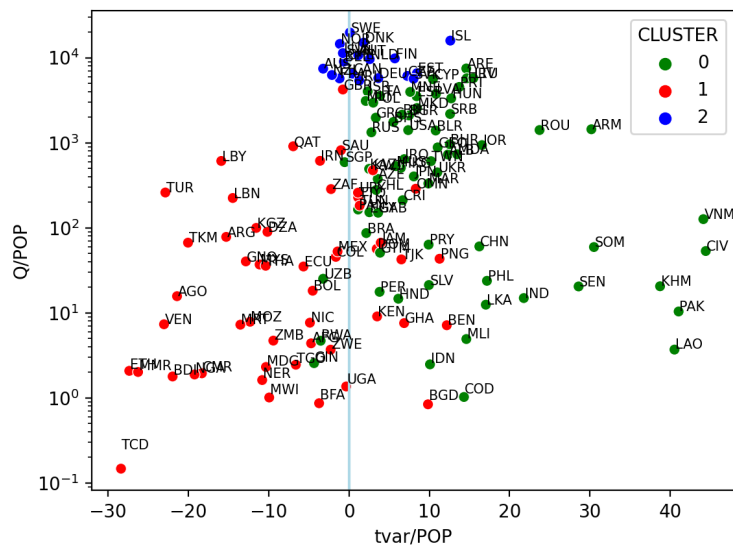


Figura J.1: Cluster Analysis in scala logaritmica

Appendice K

Grafico Cluster Crescita completo di tutti i paesi

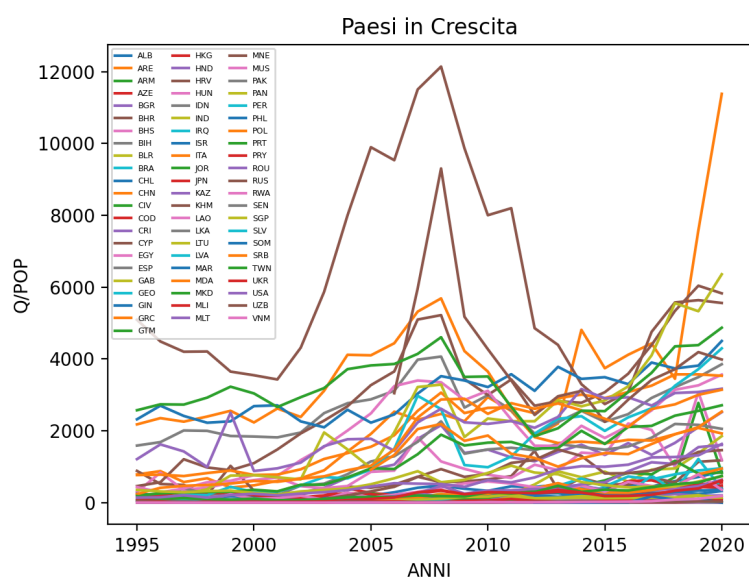


Figura K.1: Grafico cluster Crescita tutti i paesi

Appendice L

Risultati completi LDA

Tabella L.1: Paesi nello stadio di **Latenza**

Paese	Q/POP	tvar/POP	VarSt_POP	Pred.
ARG	78,1177	-15,2662	-1,0042	Latenza
BDI	1,7834	-21,9347	-1,989	Latenza
BOL	18,1832	-4,5061	-0,4373	Latenza
CHL	278,1119	3,3834	0,2974	Latenza
CMR	1,9431	-18,2911	-0,4962	Latenza
DZA	89,6007	-10,1359	-0,4337	Latenza
ECU	35,0484	-5,6525	0,7616	Latenza
ETH	2,0776	-27,3251	0,6818	Latenza
GAB	150,2323	3,634	0,3632	Latenza
GBR	4229,6721	-0,7538	-0,088	Latenza
GHA	7,5678	6,8703	-0,5429	Latenza
GTM	50,927	3,8783	0,9072	Latenza
HND	14,6434	6,15	0,2663	Latenza
IRL	5381,5109	1,2828	0,2672	Latenza
ISR	4018,6033	2,3376	0,3747	Latenza
JAM	66,7266	3,9599	-0,9713	Latenza
KEN	9,0663	3,511	-0,4189	Latenza
KGZ	99,6691	-11,5292	-1,1436	Latenza
KWT	475,3031	2,9895	-0,1696	Latenza
MDG	2,3068	-10,318	-1,7654	Latenza
MLT	3102,4115	2,0937	0,2021	Latenza
MMR	2,0147	-26,2318	-0,6425	Latenza
MOZ	7,8462	-12,2938	0,9232	Latenza
MUS	530,609	5,8387	0,2941	Latenza

Continua nella pagina successiva

Tabella L.1 – *Continua dalla pagina precedente*

Paese	Q/POP	tvar/POP	VarSt POP	Pred.
MYS	36,9839	-11,1097	-0,6141	Latenza
NER	1,619	-10,7895	-1,8115	Latenza
NGA	1,8821	-19,2015	-0,137	Latenza
NIC	7,6566	-4,8486	0,4404	Latenza
PAN	163,8636	1,1383	0,4535	Latenza
QAT	908,6864	-6,9311	-0,5437	Latenza
SAU	816,3456	-0,9946	0,1897	Latenza
TGO	2,4528	-6,6017	-0,118	Latenza
THA	35,7655	-10,3528	-0,0777	Latenza
TKM	66,8149	-20,017	-0,7552	Latenza
TTO	227,7055	1,1498	-0,6853	Latenza
UGA	1,3605	-0,3399	-0,3532	Latenza
URY	258,8654	1,1321	-0,4715	Latenza
VEN	7,3305	-23,0015	-2,0295	Latenza
ZMB	4,7091	-9,4019	-0,9583	Latenza
ZWE	3,6864	-2,276	-1,0042	Latenza
AFG	4,3787	-4,7036	-1,2069	Latenza
TCD	0,1469	-28,356	-1,3777	Latenza
GNQ	40,2255	-12,822	-1,4702	Latenza
MWI	1,0094	-9,9143	-0,9527	Latenza
ZAF	285,2188	-2,2254	-0,1327	Latenza
MRT	7,2488	-13,5021	-1,5594	Latenza

Tabella L.2: Paesi nello stadio di **Crescita**

Paese	Q/POP	tvar/POP	VarSt POP	Pred.
BEN	7,1421	12,1766	-1,2024	Crescita
BHS	1755,6747	5,5323	0,4293	Crescita
BIH	2137,9529	6,6075	0,3378	Crescita
CHN	60,3756	16,2257	-0,2614	Crescita
CIV	53,2873	44,3964	1,1461	Crescita
CRI	210,9852	6,6918	0,0125	Crescita
CYP	5591,9144	10,455	0,6287	Crescita
DNK	15038,1346	1,8384	0,1228	Crescita
ESP	3522,1283	8,4212	0,7208	Crescita
EST	6573,7844	8,4691	0,4772	Crescita

Continua nella pagina successiva

Tabella L.2 – *Continua dalla pagina precedente*

Paese	Q/POP	tvar/POP	VarSt_POP	Pred.
FIN	9819,2744	5,7124	0,4164	Crescita
GEO	881,746	11,0399	0,3773	Crescita
HKG	506,5151	6,4627	-0,0599	Crescita
HRV	5737,3012	14,6653	0,8358	Crescita
HUN	3325,1455	12,7287	0,7783	Crescita
IDN	2,4763	10,0891	0,442	Crescita
IRQ	641,3931	6,8828	0,91	Crescita
ISL	15885,6345	12,6166	0,0843	Crescita
ITA	3559,1943	3,8725	0,187	Crescita
JOR	937,8959	16,5501	-0,1336	Crescita
KHM	20,5286	38,7109	4,3473	Crescita
LKA	12,4573	17,0348	-0,2745	Crescita
LVA	3756,7816	10,833	1,1647	Crescita
MAR	331,8565	9,8858	0,9659	Crescita
MLI	4,9141	14,6103	-0,1403	Crescita
NPL	7,4996	56,6137	2,3569	Crescita
OMN	287,88	8,3159	-0,5146	Crescita
PAK	10,3639	41,0208	-0,8627	Crescita
PHL	23,7771	17,1739	-0,8208	Crescita
PNG	43,2235	11,2788	-1,6794	Crescita
PRT	4537,2252	13,7356	0,6583	Crescita
PRY	63,3508	9,9163	0,0727	Crescita
ROU	1410,5229	23,717	0,7917	Crescita
SEN	20,4228	28,5758	3,1593	Crescita
SLV	21,2156	9,9553	0,6644	Crescita
TWN	607,438	10,2112	0,9608	Crescita
USA	1407,7027	7,378	0,422	Crescita
VNM	126,3544	44,0994	2,222	Crescita
LAO	3,6941	40,5041	-0,1167	Crescita
SOM	59,3497	30,4867	-1,2589	Crescita
SRB	2180,8608	12,5925	1,0916	Crescita
MNE	3965,1188	7,6275	0,5582	Crescita

Tabella L.3: Paesi nello stadio di **Crescita**

Paese	Q/POP	tvar/POP	VarSt_POP	Pred.
BHR	960,1013	12,5551	2,4168	Saturazione
CHE	9541,0174	-0,3276	0,1488	Saturazione
EGY	152,3137	2,5477	1,1432	Saturazione
GIN	2,5579	-4,3346	1,1619	Saturazione
KAZ	493,3368	2,5174	1,6479	Saturazione
KOR	449,405	66,9772	13,0569	Saturazione
MDA	732,0127	13,4686	3,1316	Saturazione
NOR	14524,0697	-1,1282	0,1687	Saturazione
SGP	591,0173	-0,5313	1,2731	Saturazione
TZA	18,9619	13,4583	12,3503	Saturazione
UKR	448,2866	10,9955	2,7354	Saturazione
UZB	25,1453	-3,2044	3,6083	Saturazione
ARM	1440,602	30,1641	4,1891	Saturazione
RWA	4,693	-3,5125	1,3357	Saturazione
SLE	5,5441	2,0769	6,1599	Saturazione
COD	1,0239	14,3139	1,9951	Saturazione

Note

1. <https://www.exportplanning.com/it/magazine/article/2021/11/30/stadi-di-sviluppo-delle-importazioni/>
2. <https://www.exportplanning.com/it/magazine/article/2021/11/30/stadi-di-sviluppo-delle-importazioni/>
3. <https://www.exportplanning.com/it/magazine/article/2022/01/05/cogli-lattimo/>
4. W.W. Rostow, *The Economic History Review*, Second Series, Vol. XII, No.1, 1959
5. W.W. Rostow, *The Economic History Review*, Second Series, Vol. XII, No.1, pag. 1, 1959
6. W.W. Rostow, *The Economic History Review*, Second Series, Vol. XII, No.1, 1959
7. W.W. Rostow, *The Process of Economic Growth*, Chapter IV, Oxford, 1953
8. W.W. Rostow, *The Process of Economic Growth*, Chapter IV, Oxford, 1953
9. W.W. Rostow, *The Economic History Review*, Second Series, Vol. XII, No. I, pag. 4-5
10. W.W. Rostow, *The Takeoff into Self-Sustained Growth*, *The Economic Journal*, Volume 66, 1956, pages 25-48
11. W.W. Rostow, *The Economic History Review*, Second Series, Vol. XII, No. I, pag. 7, 1959
12. W.W. Rostow, *The Economic History Review*, Second Series, Vol. XII, No. I, pag 11-13, 1959
13. Industry Life Cycle, Inc.Com
14. Hitt, Michael A., R. Duane Ireland and Robert E. Hosikisson, *Strategic Management: Competitiveness and Globalization Fourth Edition*. South-Western College Publishing, 2001
15. Svend Hollensen, *Global Marketing*, Pearson, 8th Edition, pp 459-467
16. Louis T. Wells Jr, *A Product Life Cycle for International Trade?*, *Journal of Marketing*, Vol. 32, pp 1-6
17. Louis T. Wells Jr, *A Product Life Cycle for International Trade?*, *Journal of Marketing*, Vol. 32, pp 1-6
18. Raymond Vernon, *International Investment and International Trade in the Product Cycle*, *The Quarterly Journal of Economics*, 1966, Vol.80, No.1, pp. 190-207

19. Louis T. Wells Jr, A Product Life Cycle for International Trade?, Journal of Marketing, Vol. 32, pp 1-6
20. Raymond Vernon, International Investment and International Trade in the Product Cycle, The Quarterly Journal of Economics, 1966, Vol.80, No.1, pp. 190-207
21. Associazione Nazionale Enciclopedia della Banca e della Borsa, bankpedia.org
22. Borsa Italiana, <https://www.borsaitaliana.it/notizie/sotto-la-lente/cagr-259.html>
23. S. Raschka e V. Mirjalili, Machine Learning con Python, 2020, pp 2-15
24. S. Raschka e V. Mirjalili, Machine Learning con Python, 2020, pp 302-330
25. S. Raschka e V. Mirjalili, Machine Learning con Python, 2020, pp 47-94
26. S. Raschka e V. Mirjalili, Machine Learning con Python, 2020, pp 47-94
27. S. Raschka e V. Mirjalili, Machine Learning con Python, 2020, pp 140-150