

Analisi dei settori di occupazione nei principali Paesi europei

Alessandro La Farciola

21 novembre 2021

Indice

1	Introduzione	1
1.1	Contesto e prospettiva	1
1.2	Presentazione del <i>dataset</i>	1
1.3	Link alle tabelle di dati	2
2	Analisi	2
3	Conclusioni	8
4	Appendice	9

1 Introduzione

1.1 Contesto e prospettiva

L'obiettivo della nostra ricerca è quello di capire la distribuzione del numero di occupati nei vari settori economico-lavorativi a livello europeo. In particolare, sono stati acquisiti i dati sul numero di occupati di 35 Paesi europei divisi in attività economiche e settori lavorativi. A partire da questi, ci viene chiesto di capire in quale macro area lavorativa ciascun Paese impiega una maggiore quantità di forza lavoro. Un'analisi di questo tipo può risultare interessante al fine di indagare quali fattori possano influenzare le aree economico-occupazionali delle nazioni europee e se ci sono alcuni Stati che spiccano particolarmente in una direzione. L'anno considerato per la nostra indagine è il 2018, così da non avere dati influenzati da fattori esterni, come la pandemia che ha determinato un mutamento dello scenario a partire da fine 2019.

1.2 Presentazione del *dataset*

La tabella di dati da analizzare ha come osservazioni i principali Stati europei, mentre come fattori presenta il numero di occupati per attività economica relativi all'anno 2018. In particolare, le attività economiche prese in considerazione sono:

- *AgrForFish*: settore agricolo, forestale e pesca;
- *Manufacturing*: settore manifatturiero;
- *ElecGas*: elettricità, gas e vapore;
- *Construction*: settore edilizio;
- *TranspStorage*: trasporto e stoccaggio;
- *AccomFood*: settore alberghiero e di ristorazione;
- *InfoComm*: attività di informazione e comunicazione;
- *FinancInsur*: attività finanziaria e assicurativa;
- *RealEstate*: beni immobiliari;
- *ProfScientTechnical*: professioni tecniche e scientifiche;
- *AdmPubDefence*: amministrazione pubblica e forze dell'ordine;
- *Education*: settore di educazione e formazione;
- *Health*: servizi sanitari e sociali;
- *Arts*: settore artistico e di intrattenimento;
- *Water*: approvvigionamento idrico, gestione rifiuti;
- *Trade*: settore commerciale all'ingrosso e al dettaglio;
- *Employed*: numero totale di occupati.

A partire dalle tabelle sopra citate, si è provveduti ad apportare alcune modifiche ai dati forniti così da rendere l'analisi maggiormente efficiente al nostro scopo. Per i dettagli, guardare l'Appendice.

1.3 Link alle tabelle di dati

I dati selezionati per la nostra analisi sono stati interamente reperiti sul sito di Eurostat. In particolare, è possibile accedere direttamente ad essi attraverso il link: https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=lfsq_egdn2&lang=en. Da qui è possibile accedere a tutte le colonne della nostra tabella selezionando:

- *Time*: 2018Q1, 2018Q2, 2018Q3, 2018Q4;
- *Duration*: Total;
- *Sex*: Total;
- *Age class*: from 15 to 64;
- *Unit*: Thousands.

Infine, la tabella relativa al numero totale degli occupati dei vari Paesi europei presi in considerazione, relativa all'anno 2018, si trova al seguente link: https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=lfsi_emp_a_h&lang=en.

2 Analisi

In primo luogo, è opportuno analizzare la matrice delle correlazioni, al fine di incominciare a capire come si legano tra di loro i vari fattori della nostra tabella di dati.

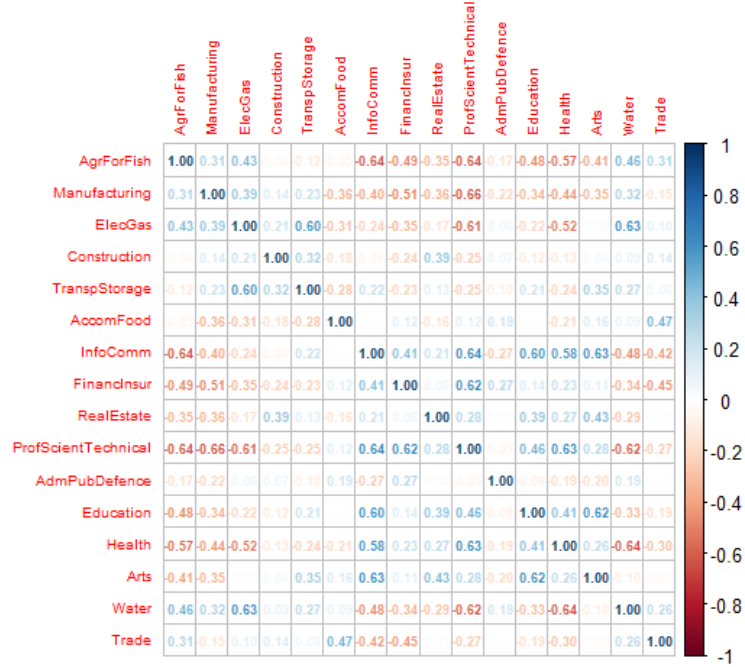


Figura 1: Matrice delle correlazioni

Osservando la figura, emerge fin da subito l'assenza di forti correlazioni tra i fattori in analisi. Nonostante ciò, è possibile notare alcune correlazioni comunque rilevanti, come ad esempio *InfoComm* oppure *ProfScientTechnical*, le quali sono ben correlate ad un buon numero di attività economiche. Tale informazione, è coerente con quanto ci saremmo aspettati a-priori, essendo l'ambito dell'informazione e l'insieme delle professioni scientifiche e tecniche legate a più settori e rami del mondo del lavoro.

Questa prima indagine ci suggerisce che i dati raccolti si prestano bene ad essere analizzati attraverso un'analisi delle componenti principali, per due ragioni importanti. Da un lato, non vi sono fattori fortemente scorrelati dagli altri che potrebbero deviare i risultati facendo emergere evidenti anomalie. Dall'altro, non potendo individuare due o più fattori fortemente correlati tra loro, ciascuno di essi fornisce informazioni determinanti al fine dell'analisi da effettuare.

Forti di tali considerazioni, procedendo con l'analisi delle componenti principali sulla tabella di dati opportunamente standardizzata (evitando così distorsioni dovute alle differenze di scale tra le diverse attività economiche considerate), è possibile individuare cinque componenti significative. Il grafico successivo (Figura 2), mostra l'andamento della varianza spiegata cumulata da cui si è ritenuto sufficiente considerare rilevanti le prime cinque componenti principali le quali raggiungono circa il 78% di varianza spiegata. Sebbene tale dato non sia particolarmente alto, si è preferito rinunciare al 5% in più di varianza spiegata dalla componente successiva al fine di non complicare ulteriormente l'analisi, rendendola meno

efficace e maneggevole.

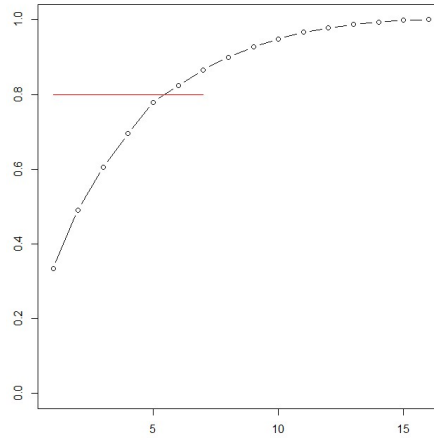


Figura 2: Varianza spiegata cumulata

A questo punto, il nostro obiettivo è quello di capire a quali componenti si allineano i vari fattori, cercando di individuare il fenomeno rappresentato dalle componenti principali selezionate. Per fare ciò, sfruttiamo i valori quantitativi forniti dalla matrice dei *loadings* rappresentata dalla seguente tabella.

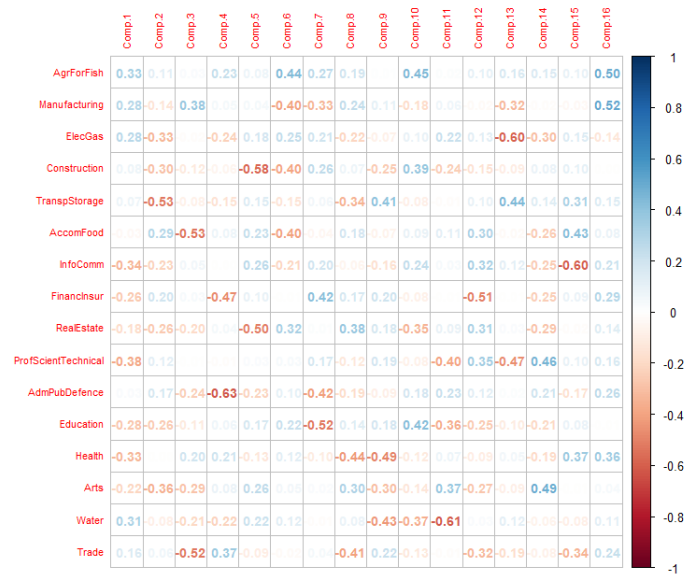


Figura 3: Matrice dei *loadings*

Alla luce dei precedenti risultati, è possibile fare le seguenti associazioni:

- 1^a componente: *AgrForFish*, *ProfScientTechnical*;

- 2^a componente: *TranspStorage*;
- 3^a componente: *AccomFood*, *Trade*;
- 4^a componente: *FinancInsur*, *AdmPubDefence*;
- 5^a componente: *Construction*, *RealEstate*.

Risulta, però, necessario ricorrere ad una rotazione delle prime cinque componenti per l'assegnamento definitivo di alcuni dei fattori, in particolare *Manufacturing*, *ElecGas*, *InfoComm*, *Education*, *Health*, *Arts*, *Water*.

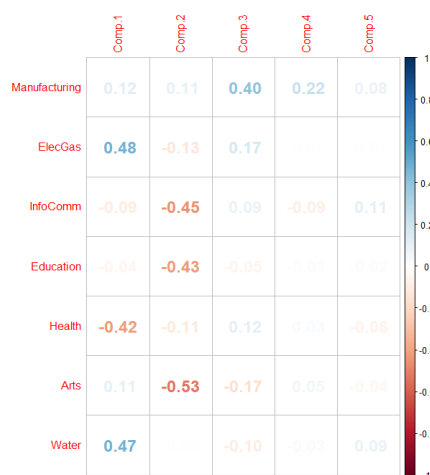


Figura 4

Come si evince dall'ultima tabella, è possibile assegnare alla prima componente i fattori *ElecGas*, *Health*, *Water*, mentre alla seconda *InfoComm*, *Education* e *Arts*, e infine *Manufacturing* alla terza. Pertanto, le assegnazioni definitive che possiamo fare sono:

- 1^a componente: *AgrForFish*, *ProfScientTechnical*, *ElecGas*, *Health*, *Water*;
- 2^a componente: *TranspStorage*, *InfoComm*, *Education*, *Arts*;
- 3^a componente: *Manufacturing*, *AccomFood*, *Trade*;
- 4^a componente: *FinancInsur*, *AdmPubDefence*;
- 5^a componente: *Construction*, *RealEstate*.

Prima di provare ad interpretare a cosa fanno riferimento le componenti principali analizzate, risulta opportuno osservare il grafico che rappresenta il piano principale.

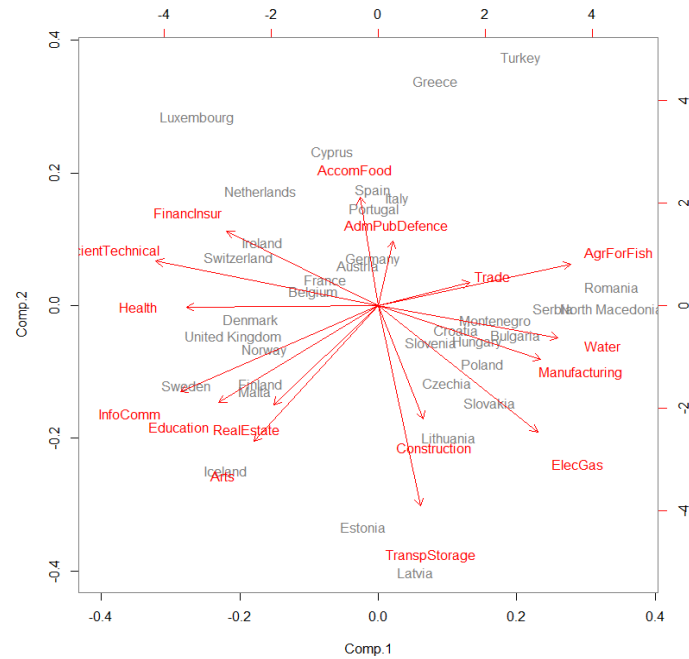


Figura 5: Piano principale

La figura mostra una netta divisione tra le nostre osservazioni. A meno di una piccola rotazione, i fattori che occupano la parte di piano con ascissa positiva rappresentano una tipologia di lavoro ben diversa da quelle che hanno la prima componente negativa. Da una parte (ascissa positiva) è possibile trovare un insieme di lavori principalmente legati al settore primario e secondario, quali ad esempio l'ambito agricolo, manifatturiero, edilizio, industriale, risorse e fonti energetiche. Dall'altra parte, invece, si raggruppano le principali tipologie di lavoro legate al settore terziario e dunque ai servizi tipici delle società più sviluppate, con particolare riferimento ai cosiddetti lavori "intellettuali", come gli ambiti legati alle attività finanziarie, quelle a tecnologia avanzata, come le telecomunicazioni o i centri di ricerca scientifica, il campo amministrativo e quello socio-educativo. Questa netta separazione può portare ad una sorta di interpretazione riassuntiva sulla tipologia di lavoro della prima componente principale. Ovvero, i Paesi che si trovano sulla parte sinistra del piano principale sono quei Paesi che impiegano un maggior numero di occupati nell'ambito della sfera dei servizi al cittadino e alle imprese, mentre dal lato opposto si concentra un'occupazione principalmente legata ai tipici lavori dei settori primario e secondario. Tale divisione è senz'altro legata all'industrializzazione e allo sviluppo economico-sociale dei Paesi considerati. Inoltre, questa componente può spiegare in un certo modo anche alcune correlazioni che si potevano osservare nella matrice delle correlazioni di apparente difficile interpretazione, come la buona correlazione tra *Water*, *ElecGas* e *AgrForFish* con *Health*.

Detto ciò, si può passare a considerare le componenti principali successive. Alla seconda si possono associare alcuni dei principali servizi, mentre per quanto riguarda la terza com-

ponente, questa assume un forte significato produttivo e commerciale. Guardando il grafico che mostra il piano tra la seconda e la terza componente (Figura 6a), si nota che il settore manifatturiero punta nella direzione positiva, mentre il settore commerciale (all'ingrosso e al dettaglio) e quello della ristorazione-hôtellerie puntano nel verso opposto. Può essere corretto ritenere che quest'ultima direzione sia influenzata in parte dal settore del turismo, in cui l'albergo, la ristorazione e il commercio ricoprono ruoli di primo piano.

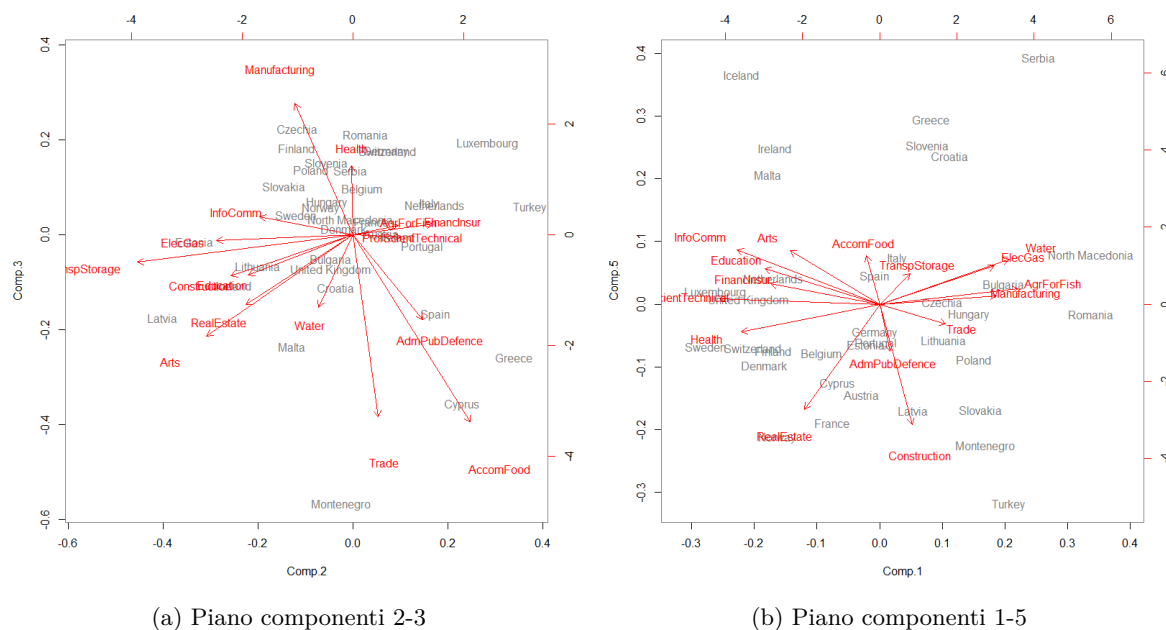


Figura 6

La quarta componente racchiude l'ambito amministrativo, ma soprattutto economico-finanziario e, infine, vi è la quinta componente che raggruppa l'intero settore immobiliare, dall'edilizia ai beni immobili. Per quest'ultima, il verso di riferimento è quello negativo e quindi, osservando la Figura 6b, si può notare una buona differenziazione tra i vari Paesi europei e, in particolare, nella parte alta emerge un gruppo di nazioni poco devote a tale settore.

isole rocciose la cui conformazione geografica impedisce un fervido sviluppo del settore delle costruzioni. Al contrario, i Paesi europei dominanti, quali la *Francia*, la *Germania*, L'*Gran Bretagna* e così via si collocano quasi sempre nella parte centrale dei grafici analizzati e questo può essere specchio dello sviluppo più completo ed equilibrato dei suddetti Paesi.

Un'ulteriore osservazione che è possibile sottolineare è la presenza del settore sanitario nella prima componente. Rappresentando, infatti, una percentuale importante in numero di occupati, esso assume un ruolo fondamentale nell'assegnare un significato riassuntivo alla prima componente, più di settori quali quello artistico o educativo che, pur avendo un valore non banale sulla prima componente (vedi matrice dei *loadings*), influenzano maggiormente il significato assegnato alla seconda componente. Come accade con *Health*, anche *Arts*, *Education* e altri fattori, pur allineandosi principalmente con una delle componenti, assumono un ruolo importante anche nella spiegazione del fenomeno di altre componenti. Questo si evince dai valori assunti nella matrice dei *loadings*, ma anche dalla rappresentazione dei vari piani principali. Tale aspetto deriva da due cause essenziali. In primo luogo, i dati raccolti sulle varie attività economiche sono molto ampi, nel senso che in una singola attività presa in considerazione rientrano molteplici tipologie di lavori e professioni che hanno un'influenza su più macro aree di occupazione. In secondo luogo, ricordiamo che l'analisi da noi effettuata cattura circa il 78% della varianza spiegata e quindi è naturale non riuscire a cogliere la variabilità dell'intero fenomeno da indagare. Nonostante ciò, la nostra analisi si ritiene sufficiente ai fini di una ricerca più superficiale sulla tipologia di occupazione dei vari Paesi europei, facendo particolare riferimento al piano principale. Per una futura indagine più accurata, invece, sarebbe opportuno considerare di suddividere alcuni dei fattori, quali *ProfScientTechnical*, *TranspStorage*, *Trade*, in più specifici settori lavorativi, così da fornire informazioni distintive per una più facile e corretta assegnazione alle componenti principali.

4 Appendice

In questa appendice, si vuole esplicitare le modifiche effettuate alla tabella di dati reperite online al fine di rendere più efficiente l'analisi effettuata.

Innanzitutto, le tabelle presentavano i dati relativi a periodi quadrimestrali e pertanto si è considerata la media aritmetica dei quattro valori al fine di possedere un unico dato annuale. Tale operazione è stata effettuata su tutti i fattori relativi alle diverse attività economiche.

In secondo luogo, si è provveduto a normalizzare il numero di occupati di ciascuna attività economica per il numero totale di impiegati nel Paese di riferimento. Tale operazione punta a non condizionare l'analisi con le differenze di estensione e, in particolare, di popolazione tra i vari Stati europei. A questo punto l'analisi è stata effettuata su tutte le colonne eccezion fatta per quella relativa al numero totale di impiegati, che è stata opportunamente eliminata.

Infine, alcuni dati forniti dal sito *Eurostat* non erano completi. Quindi per le osservazioni che per un determinato fattore non presentavano il dato raccolto per ciascun quadrimestre, si è considerata la media aritmetica tra quelle disponibili. Questo è il caso dell'*Estonia* e dell'*Islanda* per la colonna *Water*; degli Stati *Croazia*, *Lussemburgo*, *Islanda*, *Montenegro*, *Macedonia del Nord* e *Serbia* per la colonna *RealEstate*. D'altro canto, nei casi di *Lussemburgo*, *Malta* per la colonna *ElecGas* e del solo *Lussemburgo* per la colonna *Water* il dato

era mancante per l'intero anno. In questi ultimi casi si sarebbe potuto procedere inserendo manualmente come dato la media della colonna presa in considerazione, così da condizionare il meno possibile la varianza del fattore e quindi l'analisi in generale. Con questa scelta, però, si veniva a creare una forte anomalia, dato che la media aritmetica tra gli altri Paesi europei forniva un dato eccessivamente elevato rispetto all'estensione demografica dei Paesi quali *Lussemburgo* e *Malta*. Pertanto, si è preferito inserire manualmente un dato che fosse più coerente con il significato delle osservazioni catturando tale informazione da quelle che, in base alle scale dei dati, risultavano più simili, come *Estonia*, *Cipro* e *Montenegro*.