ePublishing
ΕΘΝΙΚΟ ΚΕΝΤΡΟ ΤΕΚΜΗΡΙΩΣΗΣ
NATIONAL DOCUMENTATION CENTRE

Τετράδια Ανάλυσης Δεδομένων

Τόμ. 20, Αρ. 1 (2024)

Τετράδια Ανάλυσης Δεδομένων - 20

Το στην ερμηνεία του παραγοντικού επιπέδου με συμπληρωματικές μεταβλητές: Εφαρμογή στην Αναλυτική χρηματοοικονομικών δεδομένων

Στράτος Μοσχίδης, Αθανάσιος Κ. Θανόπουλος

Copyright © 2024, Τετράδια Ανάλυσης Δεδομένων

Άδεια χρήσης Creative Commons Attribution-NonCommercial-NoDerivatives 4.0.

Βιβλιογραφική αναφορά:

Μοσχίδης Σ., & Θανόπουλος Α. Κ. (2024). Το <<ερμηνευτικό επίπεδο>> στην ερμηνεία του παραγοντικού επιπέδου με συμπληρωματικές μεταβλητές: Εφαρμογή στην Αναλυτική χρηματοοικονομικών δεδομένων . *Τετράδια Ανάλυσης Δεδομένων*, *20*(1), 24–31. ανακτήθηκε από https://ejournals.epublishing.ekt.gr/index.php/dab/article/view/32333

Το <<ερμηνευτικό επίπεδο>> στην ερμηνεία του παραγοντικού επιπέδου με συμπληρωματικές μεταβλητές: Εφαρμογή στην Αναλυτική χρηματοοικονομικών δεδομένων

Στράτος Μοσχίδης[1], Αθανάσιος Θανόπουλος[1]

[1]Ελληνική Στατιστική Αρχή

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ

Αναλυτική δεδομένων
Μηχανική μάθηση
Λογιστική- Χρηματοοικονομική
Διερευνητική ανάλυση
Πολυμεταβλητή οπτικοποίηση

ΣΤΟΙΧΕΙΑ ΕΠΙΚΟΙΝΩΝΙΑΣ

Στράτος Μοσχίδης
Ελληνική Στατιστική Αρχή
Email: smos@uom.edu.gr

ΠΕΡΙΛΗΨΗ

Η παραγοντική μέθοδος των πολλαπλών αντιστοιχιών(MCA) είναι μέθοδος της μη επιβλεπόμενης μηχανικής μάθησης που εφαρμόζεται σε κατηγορικής φύσεως δεδομένα προκειμένου να διερευνηθούν οι εντονότερες αλληλοεπιδράσεις μεταξύ των κατηγοριών τους, καθώς και οι κατευθύνσεις(τάσεις) της διασποράς τους. Οι συμπληρωματικές μεταβλητές που χρησιμοποιούνται και στην(MCA) δεν συμμετέχουν στην κατασκευή των παραγοντικών αξόνων, απλώς προβάλλονται μέσω των σχέσεων μεταφοράς μαζί με τα υπόλοιπα σημεία (κατηγορίες μεταβλητών) στα παραγοντικά διαγράμματα και μπορούν να συμβάλουν περαιτέρω στην ερμηνεία του φαινομένου. Στην παρούσα εργασία εξετάζεται ο τρόπος με τον οποίο μπορεί να ενσωματωθεί η γραφική πληροφορία από το «Ερμηνευτικό Επίπεδο» στα παραγοντικά επίπεδα που περιλαμβάνουν συμπληρωματικά σημεία. Μέσω του Ερμηνευτικού Επιπέδου εντοπίζονται τα σημαντικά ερμηνευτικά σημεία του παραγοντικού επιπέδου και στη συνέχεια εκείνα από αυτά που είναι τα «κοντινά» σε συμπληρωματικά σημεία. Έτσι οι χρήστες που επιθυμούν να χρησιμοποιήσουν συμπληρωματικές μεταβλητές στις αναλύσεις τους μπορούν να πετύχουν καλύτερα και ασφαλέστερα ερμηνευτικά αποτελέσματα-συμπεράσματα. Τα ως άνω εφαρμόστηκαν στη διερεύνηση του συστήματος ανταμοιβών, στην επίδοση, ανάπτυξη και αποτελεσματικότητα των επιχειρήσεων. Οι μεταβλητές τομέας επιχείρησης και η ιεραρχική θέση των εργαζομένων στην επιχείρηση ελήφθησαν ως συμπληρωματικές.

Εισαγωγή

Η παραγοντική ανάλυση των πολλαπλών αντιστοιχιών είναι μια μέθοδος της μη επιβλεπόμενης μάθησης η οποία χρησιμοποιείται για την ανακάλυψη σχέσεων και αλληλεξαρτήσεων μεταξύ των κατηγοριών ενός σετ κατηγορικών μεταβλητών (Fithian & Josse, 2017; Greenacre & Blasius, 2006). Η μέθοδος μπορεί να χαρακτηριστεί ως μια γενίκευση της ανάλυσης σε κύριες συνιστώσες με την διαφορά ότι οι μεταβλητές που μελετώνται είναι κατηγορικής φύσης(Abdi, 2007). Ένα ενδιαφέρον σημείο της ανάλυσης των πολλαπλών αντιστοιχιών αποτελούν οι λεγόμενες συμπληρωματικές μεταβλητές. Πρόκειται για μεταβλητές οι οποίες δεν συμμετέχουν στον σχηματισμό των παραγοντικών αξόνων (Graffelman & Aluja-Banet, 2003), δηλαδή δεν έχουν μάζα, όμως μπορούν να προβληθούν στους παραγοντικούς χάρτες (Torres & van de Velden, 2007) και να συμμετέχουν στην ερμηνεία τους. Οι υπολογισμοί για αυτά τα σημεία έχουν παρουσιαστεί από τον (Gabriel, 1995). Επιπροσθέτως, ένα βιβλιογραφικά καταγεγραμμένο πρόβλημα είναι αυτό της ασφαλούς ερμηνείας των σημαντικότερων σημείων στα παραγοντικά διαγράμματα καθώς έχει απασχολήσει διάφορους ερευνητές(Greenacre, 1993; Hoffman & De Leeuw, 1992) . Μια ασφαλής ερμηνεία των αποτελεσμάτων της μεθόδου εκτός της σχετικής θέσης των σημείων μέσα από τις συντεταγμένες τους στα παραγοντικά διαγράμματα περιλαμβάνει την συν-αξιολόγηση επιμέρους δεικτών όπως η συνεισφορά και η ποιότητα προβολής (Greenacre, 1991). Αυτή η εργασία μολονότι είναι αποτελεσματική συχνά μπορεί να οδηγήσει σε εσφαλμένες ερμηνείες. Το

© 2024, Συγγραφείς
Άδεια CC-BY-SA 4.0

φαινόμενο γίνεται ακόμα πιο έντονο κατά την ερμηνεία των σημαντικότερων σημείων ενός παραγοντικού επιπέδου. Στην περίπτωση αυτή επιπλέον υπολογισμοί απαιτούνται προκειμένου να υπολογιστούν δείκτες όπως είναι η συνεισφορά και η ποιότητα προβολής ενός σημείου. Ειδικότερα για τους μη ειδικούς χρήστες της μεθόδου, οι επιπλέον υπολογισμοί αποτελούν μια επίπονη εργασία. Το παραπάνω πρόβλημα αντιμετωπίζεται με την πρόταση της έννοιας της ερμηνευτικής συντεταγμένης η οποία συνιστά την βάση για την κατασκευή των αντίστοιχων ερμηνευτικών αξόνων και επιπέδων. Σύμφωνα με αυτή την προσέγγιση κάθε σημείο διατηρεί τον προσανατολισμό του (δεξιά ή αριστερά) από την αρχή ενός παραγοντικού άξονα ενώ η συντεταγμένη του ορίζεται ως η απόλυτη αδράνειά του που συνεισφέρει στον συγκεκριμένο άξονα. Με αυτόν τον τρόπο εξασφαλίζεται ενσωματώνεται η πληροφορία της κατεύθυνσης από τις κλασικές συντεταγμένες και η συνεισφορά του σημείου προς τον άξονα (Moschidis et al., 2022). Το αποτέλεσμα είναι η δημιουργία ενός άξονα ο οποίος περιλαμβάνει σημεία που προβάλλονται στον παραγοντικό άξονα όμως με διάταξη σημαντικότητας ως προς την συνεισφορά τους στην δημιουργία του αντίστοιχου παραγοντικού άξονα. Η έννοια του ερμηνευτικού άξονα γενικεύεται στο παραγοντικό επίπεδο όπου αποδεικνύεται ότι κάθε ένα σημείο του ερμηνευτικού επιπέδου βρίσκεται επάνω σε έναν γεωμετρικό τόπο σημείων (τετράγωνο) όπου όλα τα σημεία τα οποία βρίσκονται στο ίδιο τετράγωνο έχουν την ίδια συνεισφορά στην δημιουργία του επιπέδου. Με τον τρόπο αυτό η σύγκριση διαφορετικών σημείων στο παραγοντικό επίπεδο γίνεται σημαντικά ευκολότερη, γρηγορότερη και ασφαλέστερη. Δυστυχώς ένας από τους περιορισμούς του ερμηνευτικού επιπέδου είναι ότι δεν μπορεί να συμπεριλάβει συμπληρωματικά σημεία καθώς αυτά όπως προαναφέρθηκε δεν έχουν μάζα συνεπώς δεν μπορεί να προχωρήσει ο αντίστοιχος υπολογισμός για την εξαγωγή των ερμηνευτικών τους συντεταγμένων. Όμως στην παρούσα εργασία θα δώσουμε μια εναλλακτική προσέγγιση που χειρίζεται συνδυαστικά τις πληροφορίες από το ερμηνευτικό επίπεδο και τις ενσωματώνει στο κλασικό παραγοντικό επίπεδο. Στόχος συνεπώς αυτή της εργασίας είναι ένας αυτόματος προσδιορισμός των εντονότερων-σημαντικότερων ερμηνευτικά σημείων του παραγοντικού επιπέδου που βρίσκονται και τα συμπληρωματικά σημεία.

Μεθοδολογία - Ανάλυση

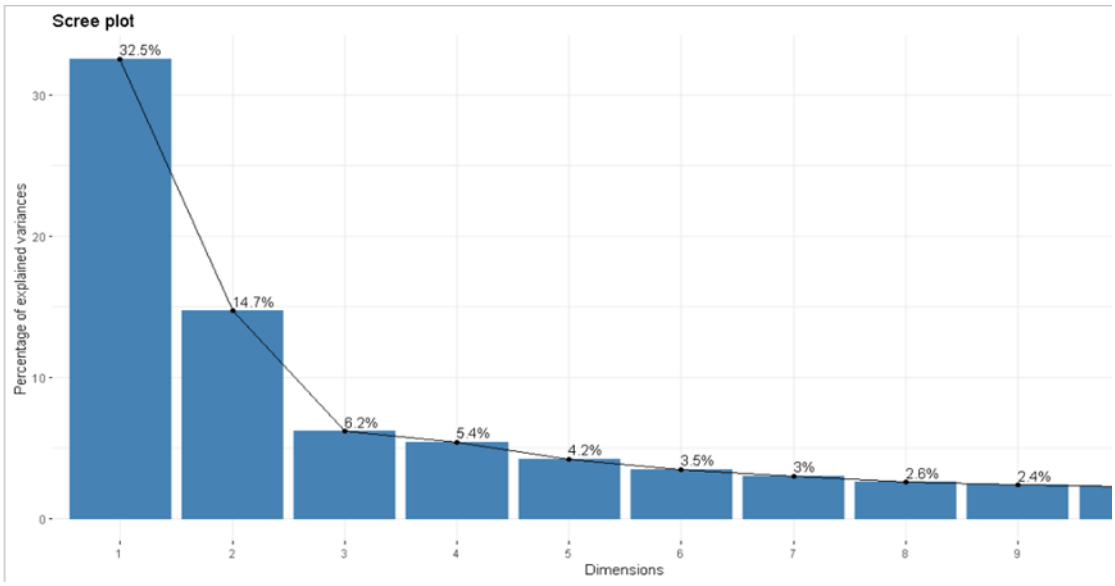
Στο σημείο αυτό ορίζεται ο τρόπος δημιουργίας ενός εναλλακτικού παραγοντικού επιπέδου στο οποίο εμφανίζονται μόνο σημεία τα οποία είναι σημαντικά για την δημιουργία του ερμηνευτικού επιπέδου και κατά συνέπεια του παραγοντικού επιπέδου. Δηλαδή, θα πραγματοποιηθεί η κατασκευή ενός παραγοντικού επιπέδου «δια μέσω» του ερμηνευτικού επιπέδου. Με τον τρόπο αυτό τα προβαλλόμενα συμπληρωματικά σημεία στο παραγοντικό άξονα και επίπεδο θα βρίσκονται δίπλα σε σημεία τα οποία θα έχουν σημαντική ερμηνευτική αξία όπως αυτό προκύπτει από τον αντίστοιχο ερμηνευτικό επίπεδο. Η προτεινόμενη λύση δίνεται μέσα από εμπειρικά δεδομένα από τον χώρο της λογιστικής και χρηματοοικονομικής. Για την εφαρμογή της μεθόδου αλλά και των οπτικοποιήσεων χρησιμοποιήθηκε η γλώσσα προγραμματισμού R μέσα από τα πακέτα FactoMiner (Lê et al., 2008) ggplot2 (Wickham et al., 2016) και Tidyverse (Wickham et al., 2019). Τα δεδομένα αποτελούν απαντήσεις ερωτηματολογίου στο οποίο απάντησαν 1050 άτομα. Το αντικείμενο του ερωτηματολογίου αποτελεί η διερευνητική εξέταση ποιοτικών δεικτών που αφορούν την ανταμοιβή εργαζομένων, την κατάρτισή τους, την επίδοσή τους και της αποτελεσματικότητας των επιχειρήσεων. Πιο κάτω ακολουθεί η περιγραφή του σετ των δεδομένων και η κωδικοποίηση των μεταβλητών.

Πίνακας 1
Κωδικοποίηση μεταβλητών

Μεταβλητή	**Κωδικοποίηση**
ΤΟΜΕΑΣ: βασικός τομέας όπου δραστηριοποιείται η επιχείρηση	ΤΟΜΕΑΣ_1=ΒΙΟΜΗΧΑΝΙΑ, ΤΟΜΕΑΣ_2=ΥΠΗΡΕΣΙΕΣ, ΤΟΜΕΑΣ_ 3=ΕΜΠΟΡΙΟ
ΕΚ2:Η επιχείρησή μου παρέχει εκπαίδευση που επικεντρώνεται σε δεξιότητες ομαδικότητας και ομαδικής εργασίας	ΕΚ2_1=ΔΙΑΦΩΝΩ ΑΠΟΛΥΤΑ,…, ΕΚ2_5=ΣΥΜΦΩΝΩ ΑΠΟΛΥΤΑ

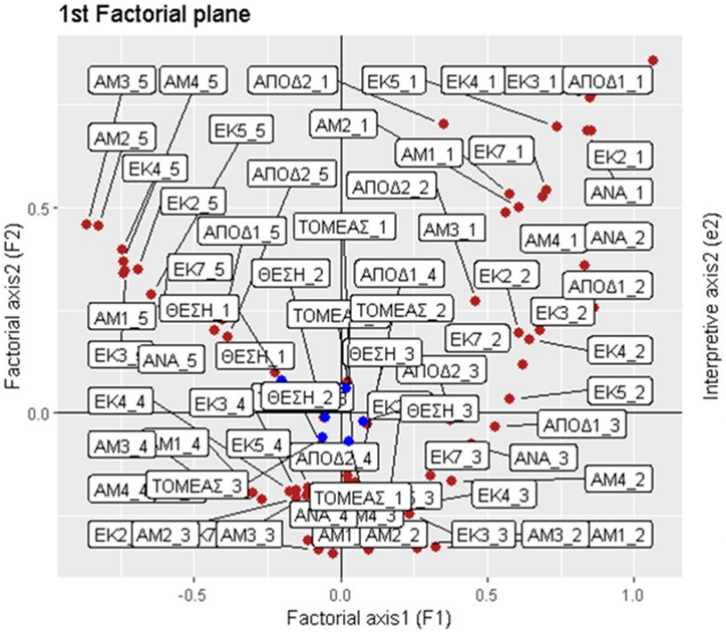
ΕΚ3:Η επιχείρησή μου παρέχει εξειδικευμένη κατάρτιση και ανάπτυξη για τους υπαλλήλους της.	ΕΚ3_1=ΔΙΑΦΩΝΩ ΑΠΟΛΥΤΑ,..., ΕΚ3_5=ΣΥΜΦΩΝΩ ΑΠΟΛΥΤΑ
ΕΚ4:Η επιχείρησή μου ξεκινάει και παρέχει διάφορα είδη εκπαίδευσης και ανάπτυξης στους υπαλλήλους της.	ΕΚ4_1=ΔΙΑΦΩΝΩ ΑΠΟΛΥΤΑ,..., ΕΚ4_5=ΣΥΜΦΩΝΩ ΑΠΟΛΥΤΑ
ΕΚ5:Παρέχονται προγράμματα τυπικής κατάρτισης για να μάθουν οι νέοι προσληφθέντες τις δεξιότητες που χρειάζονται για την εκτέλεση της εργασίας τους.	ΕΚ5_1=ΔΙΑΦΩΝΩ ΑΠΟΛΥΤΑ,..., ΕΚ5_5=ΣΥΜΦΩΝΩ ΑΠΟΛΥΤΑ
ΕΚ7:Οι εργαζόμενοι έχουν ευκαιρίες για εξέλιξη προς τα πάνω.	ΕΚ7_1=ΔΙΑΦΩΝΩ ΑΠΟΛΥΤΑ,..., ΕΚ7_5=ΣΥΜΦΩΝΩ ΑΠΟΛΥΤΑ
ΑΜ1:Μέσα στην επιχείρησή μου, τα μέλη της ομάδας ανταμείβονται με βάση την ατομική τους απόδοση πέραν του βασικού μισθού τους.	ΑΜ1_1=ΔΙΑΦΩΝΩ ΑΠΟΛΥΤΑ,..., ΑΜ1_5=ΣΥΜΦΩΝΩ ΑΠΟΛΥΤΑ
ΑΜ2:Μέσα στην επιχείρησή μου, τα μέλη της ομάδας ανταμείβονται με βάση την απόδοση της ομάδας εκτός από τον βασικό μισθό τους	ΑΜ2_1=ΔΙΑΦΩΝΩ ΑΠΟΛΥΤΑ,..., ΑΜ2_5=ΣΥΜΦΩΝΩ ΑΠΟΛΥΤΑ
ΑΜ3:Μέσα στην επιχείρησή μου, τα μέλη της ομάδας ανταμείβονται με βάση τις οργανωτικές επιδόσεις πέρα από τον βασικό μισθό τους.	ΑΜ3_1=ΔΙΑΦΩΝΩ ΑΠΟΛΥΤΑ,..., ΑΜ3_5=ΣΥΜΦΩΝΩ ΑΠΟΛΥΤΑ
ΑΜ4:Το σύστημα αμοιβών της επιχείρησής μου αντικατοπτρίζει τη συμβολή του εργαζόμενου στην επιχείρηση.	ΑΜ4_1=ΔΙΑΦΩΝΩ ΑΠΟΛΥΤΑ,..., ΑΜ4_5=ΣΥΜΦΩΝΩ ΑΠΟΛΥΤΑ
ΑΠΟΔ1:Αποτελεσματικότητα (εάν η επιχείρηση επιτυγχάνει τους στόχους της)	ΑΠΟΔ1_1=ΠΟΛΥ ΚΑΚΗ,..., ΑΠΟΔ1_5=ΠΟΛΥ ΚΑΛΗ
ΑΠΟΔ2:Αποδοτικότητα (εάν η επιχείρηση χρησιμοποιεί τα λιγότερα δυνατά διαθέσιμα για την επίτευξη των στόχων της).	ΑΠΟΔ2_1=ΠΟΛΥ ΚΑΚΗ,..., ΑΠΟΔ2_5=ΠΟΛΥ ΚΑΛΗ
ΑΝΑ:Ανάπτυξη (εάν η επιχείρηση αναπτύσσεται μέσα στις δυνατότητές της για την επίτευξη μελλοντικών ευκαιριών και προκλήσεων)	ΑΝΑ_1=ΠΟΛΥ ΚΑΚΗ,..., ΑΝΑ_5=ΠΟΛΥ ΚΑΛΗ
ΘΕΣΗ:ΘΕΣΗ ΣΤΗΝ ΙΕΡΑΡΧΙΑ	ΘΕΣΗ_1= ΑΝΩΤΑΤΟ ΣΤΕΛΕΧΟΣ, ΘΕΣΗ_2= ΜΕΣΑΙΟ ΣΤΕΛΕΧΟΣ/ΔΙΕΥΘΥΝΤΗΣ, ΘΕΣΗ_3=ΥΠΑΛΛΗΛΟΣ

Πρέπει να υπογραμμιστεί σε αυτό το σημείο ότι οι μεταβλητές ΤΟΜΕΑΣ και ΘΕΣΗ θα χρησιμοποιηθούν ως συμπληρωματικές μεταβλητές στην ανάλυση που ακολουθεί. Αναλύοντας τον πίνακα Burt ως πίνακα εισόδου στην μέθοδο των πολλαπλών αντιστοιχιών λαμβάνουμε το επόμενο διάγραμμα που συνοψίζει το ποσό της συνολικής αδράνειας των δεδομένων που ερμηνεύουν οι πρώτοι παραγοντικοί άξονες.



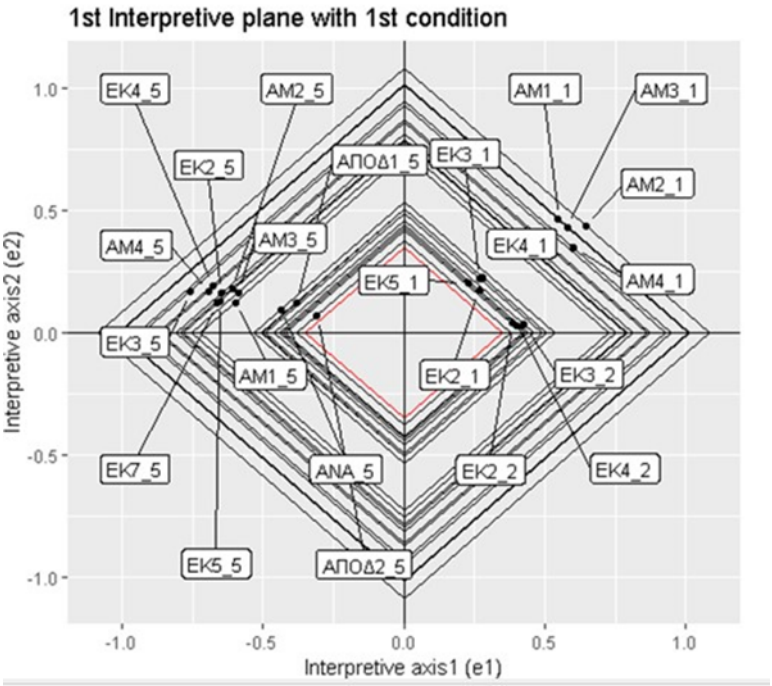
Εικόνα 1. Διάγραμμα Screeplot

Όπως φαίνεται από την παραπάνω εικόνα οι πρώτοι δύο παραγοντικοί άξονες αθροίζουν περίπου 47% της συνολικής αδράνειας του φαινομένου. Στην συνέχεια η υπόλοιπη διακύμανση επιμερίζεται με μικρές μεταβολές στους υπόλοιπους παραγοντικούς άξονες. Εμείς επικεντρώνουμε τις παρατηρήσεις μας στο πρώτο παραγοντικό επίπεδο ή εναλλακτικά στο πρώτο επίπεδο που σχηματίζεται από τις δύο πρώτες κυρίαρχες τάσεις του φαινομένου. Λαμβάνοντας τα σημεία του νέφους των στηλών στο πρώτο και δεύτερο παραγοντικό άξονα δημιουργείται το ακόλουθο επίπεδο το οποίο παρουσιάζεται στην επόμενη εικόνα.



Εικόνα 2. το πρώτο παραγοντικό επίπεδο

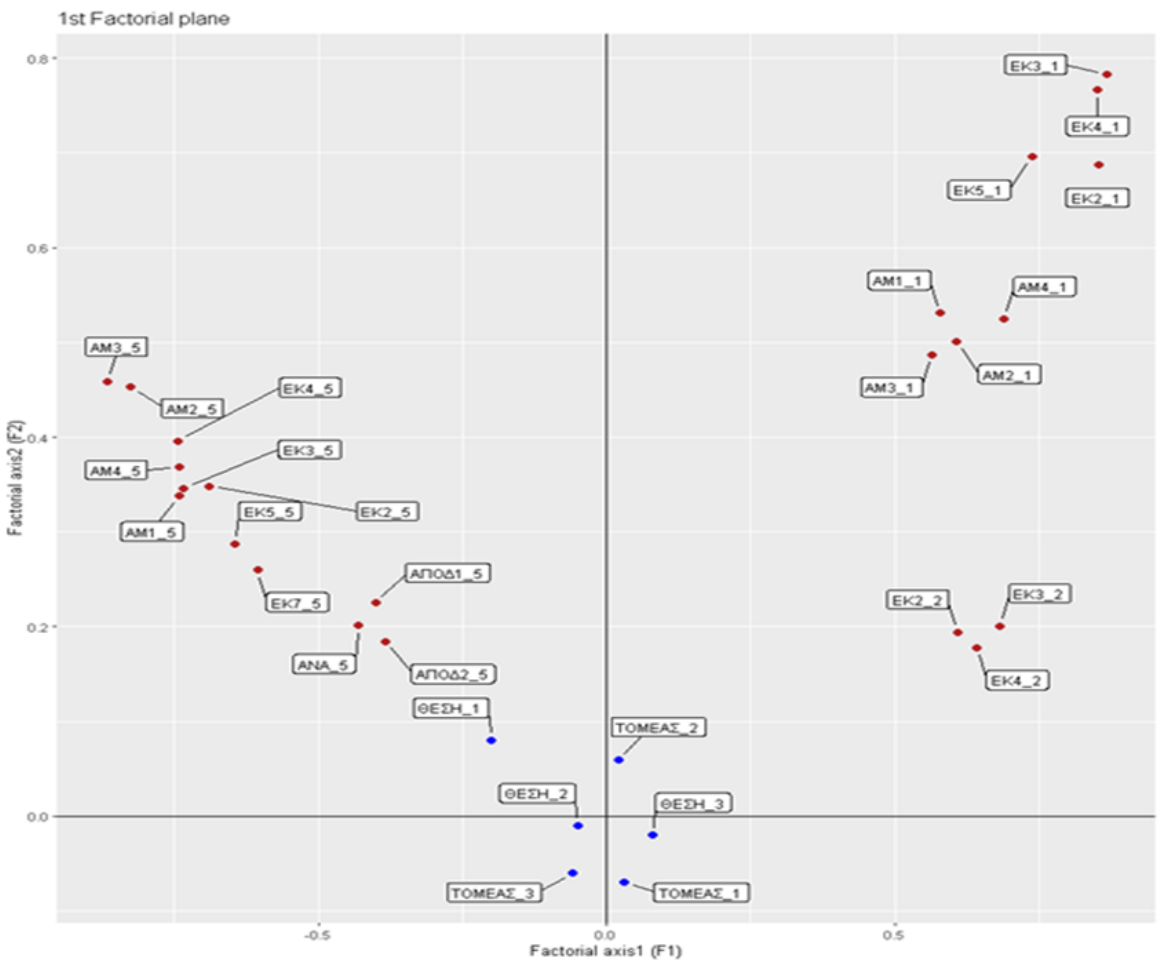
Πρέπει να σημειωθεί ότι τα σημεία με χρώμα μπλε αντιστοιχούν σε κατηγορίες των συμπληρωματικών μεταβλητών. Φυσικά από το διάγραμμα είναι δύσκολο να βγάλουμε συμπεράσματα και να ερμηνεύσουμε τις ομάδες των σημείων που σχηματίζονται σε κάθε τεταρτημόριο. Είναι λογικό να θέλουμε να διατηρήσουμε μόνο τα σημεία που είναι σημαντικότερα για το πρώτο επίπεδο. Ποια είναι όμως τα σημαντικότερα σημεία του επιπέδου; Μια λογική απάντηση θα ήταν τα σημεία εκείνα που συνεισφέρουν περισσότερο από τον μέσο όρο συνεισφοράς του πρώτου παραγοντικού επιπέδου. Αυτό μπορεί να απαντηθεί είτε με την χρήση επιπρόσθετων υπολογισμών (πράγμα το οποίο μπορεί να οδηγήσει σε σφάλματα ειδικά στην περίπτωση μη ειδικών χρηστών) είτε μέσα από το πρώτο ερμηνευτικό επίπεδο το οποίο και παρουσιάζεται στην παρακάτω εικόνα.



Εικόνα 3. τα σημαντικότερα σημεία του πρώτου ερμηνευτικού επιπέδου

Από το πρώτο ερμηνευτικό επίπεδο τεκμαίρεται ότι στο πρώτο τεταρτημόριο έχουμε τα σημεία ΕΚ2_2,ΕΚ3_2,ΕΚ4_2, ΕΚ2_1, ΕΚ3_2,ΕΚ5_1, ΑΜ1_1,ΑΜ2_1,ΑΜ3_1,ΑΜ4_1,ΕΚ4_1 ενώ στο δεύτερο τεταρτημόριο διατηρούνται ως σημαντικότερα τα σημεία ΕΚ7_5, ΕΚ5_5, ΑΜ1_5, ΑΝΑ_5, ΑΠΟΔ2_5, ΕΚ3_5, ΑΜΑ4_5ΑΜ3_5,ΑΠΟΔ1_5,ΕΛ2_5,ΕΚ4_5 ΚΑΙ ΑΜ2_5. Σύμφωνα με τον ορισμό του πρώτου ερμηνευτικού επιπέδου σημεία που βρίσκονται στην περίμετρο του ίδιου τετραγώνου έχουν ίση συνεισφορά στην δημιουργία του πρώτου παραγοντικού επιπέδου ενώ ένα σημείο που βρίσκεται σε πιο απομακρυσμένο τετράγωνο έχει μεγαλύτερη συνεισφορά από ένα άλλο που βρίσκεται σε πιο κοντινό.

Συνεπώς μόνο τα σημεία τα οποία προβάλλονται στο ερμηνευτικό επίπεδο πρέπει να προβληθούν και στο παραγοντικό επίπεδο ώστε να εξασφαλίσουμε ότι τα «πλησιέστερα» σε συμπληρωματικά σημεία είναι και σημαντικά σημεία για την δημιουργία του πρώτου παραγοντικού άξονα. Το νέο παραγοντικό επίπεδο θα είναι όπως φαίνεται στην παρακάτω εικόνα:



Εικόνα 4. το πρώτο παραγοντικό επίπεδο με συμπληρωματικές μεταβλητές και τα σημαντικότερά του σημεία.

Συνεπώς μπορούμε από το «νέο» παραγοντικό επίπεδο να ερμηνεύσουμε ότι έχουμε μια ομάδα ατόμων που οι επιχειρήσεις τους δραστηριοποιούνται στον χώρο των υπηρεσιών (ΤΟΜΕΑΣ_2), που απαντούν ότι διαφωνούν λίγο στο ότι η επιχείρησή τους παρέχει εκπαίδευσης που επικεντρώνεται σε δεξιότητες ομαδικότητας και ομαδικής εργασίας(ΕΚ3_2), που απαντούν ότι διαφωνούν λίγο στο ότι η επιχείρησή τους παρέχει εξειδικευμένη κατάρτιση και ανάπτυξη για τους υπαλλήλους της (ΕΚ3_2)και που διαφωνούν λίγο στο ότι η επιχείρησή τους ξεκινάει και παρέχει διάφορα είδη εκπαίδευσης και ανάπτυξης τα στους υπαλλήλους της(ΕΚ4_2). Επίσης μπορούμε να πούμε ότι έχουμε μια δεύτερη ομάδα ατόμων τ α οποία είναι ανώτατα στελέχη των επιχειρήσεών τους (ΘΕΣΗ_1), που απαντούν ότι οι επιχειρήσεις τους αναπτύσσονται μέσα στις δυνατότητές τους για την επίτευξη μελλοντικών ευκαιριών και προκλήσεων σε πολύ καλό βαθμό (ΑΝΑ_5), που απαντούν ότι η αποτελεσματικότητα των επιχειρήσεών τους είναι πολύ καλή και που απαντούν ότι η αποδοτικότητα των επιχειρήσεών τους είναι πολύ καλή. Κατά συνέπεια γίνεται ξεκάθαρο, ότι με την ενσωμάτωση της πληροφορίας η οποία διοχετεύεται από το ερμηνευτικό επίπεδο , η ερμηνεία του παραγοντικού επιπέδου με τις συμπληρωματικές μεταβλητές γίνεται ασφαλέστερη και ακριβέστερη.

Συμπεράσματα

Η παρούσα εργασία μελετά το πρόβλημα της σωστής ερμηνείας κατά την ανάλυση με την μέθοδο των πολλαπλών αντιστοιχιών με συμπληρωματικές μεταβλητές. Οι συμπληρωματικές μεταβλητές που συχνά χρησιμοποιούνται κατά την εφαρμογή της συγκεκριμένης μεθόδου μπορούν να εμπλουτίσουν την ερμηνεία ενός φαινομένου. Όμως ένας μη ειδικός χρήστης της μεθόδου μπορεί να οδηγηθεί σε λανθασμένη ερμηνεία του φαινομένου λόγω της συμπερίληψης μη σημαντικών σημείων προερχόμενων από τις κατηγορίες των «ενεργών» μεταβλητών. Προς την κατεύθυνση της αντιμετώπισης αυτού του ζητήματος κατασκευάστηκε νέα παραγοντικό

διάγραμμα (επίπεδο) ,στο οποίο διατηρούνται εκτός από τα σημεία των συμπληρωματικών μεταβλητών ,σημεία που είναι τα σημαντικότερα για την δημιουργία του παραγοντικού επιπέδου, όπως αυτό εξάγεται μέσα από το ερμηνευτικό επίπεδο. Κατά αυτόν τον τρόπο ο χρήστης διευκολύνεται στην σωστή ερμηνεία του φαινομένου καθώς τα συμπληρωματικά σημεία τα οποία χρησιμοποιεί βρίσκονται σε ίδιο διάγραμμα μόνο με τα σημαντικότερα σημεία.

Βιβλιογραφικές παραπομπές

Abdi, H., & Valentin, D. (2007). Multiple correspondence analysis. *Encyclopedia of Measurement and Statistics*, *2*(4), 651–657.

Fithian, W., & Josse, J. (2017). Multiple correspondence analysis and the multilogit bilinear model. *Journal of Multivariate Analysis*, *157*, 87–102. https://doi.org/10.1016/j.jmva.2017.02.009

Gabriel, K. (1995). Biplot display of multivariate categorical data, with comments on multiple correspondence analysis. *Recent Advances in Descriptive Multivariate Analysis*, *190*, 226.

Graffelman, J., & Aluja-Banet, T. (2003). Optimal Representation of Supplementary Variables in Biplots from Principal Component Analysis and Correspondence Analysis. *Biometrical Journal*, *45*(4), 491–509. https://doi.org/10.1002/bimj.200390027

Greenacre, M., & Blasius, J. (Eds.). (2006). *Multiple Correspondence Analysis and Related Methods* (0 ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781420011319

Greenacre, M. J. (1991). Interpreting multiple correspondence analysis. *Applied Stochastic Models and Data Analysis*, *7*(2), 195–210. https://doi.org/10.1002/asm.3150070208

Greenacre, M. J. (1993). Biplots in correspondence analysis. *Journal of Applied Statistics*, *20*(2), 251–269.

Hoffman, D. L., & De Leeuw, J. (1992). Interpreting multiple correspondence analysis as a multidimensional scaling method. *Marketing Letters*, *3*(3), 259–272.

Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: An R Package for Multivariate Analysis. *Journal of Statistical Software*, *25*(1). https://doi.org/10.18637/jss.v025.i01

Moschidis, S., Markos, A., & Thanopoulos, A. C. (2022). "Automatic" interpretation of multiple correspondence analysis (MCA) results for nonexpert users, using R programming. *Applied Computing and Informatics*. https://doi.org/10.1108/ACI-07-2022-0191

Torres, A., & van de Velden, M. (2007). Perceptual mapping of multiple variable batteries by plotting supplementary variables in correspondence analysis of rating data. *Food Quality and Preference*, *18*(1), 121–129. https://doi.org/10.1016/j.foodqual.2005.09.005

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, 4(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., Chang, W., & Wickham, M. H. (2016). Package 'ggplot2.' *Create Elegant Data Visualisations Using the Grammar of Graphics*. Version, 2(1), 1–189.

The <<interpretive plane>> in the interpretation of the factorial plane with supplementary variables: Application to Financial Data Analytics

Efstratios Moschidis[1], Athanasios Thanopoulos[1]

[1] Hellenic Statistical Authority

KEYWORDS

Data analytics
Machine learning
Accounting - Finance
Exploratory analysis
Multivariate visualization

CORRESPONDENCE

Efstratios Moschidis
Hellenic Statistical Authority
Email: smos@uom.edu.gr

ABSTRACT

The factorial method of multiple correspondences (MCA) is a method of unsupervised machine learning applied to categorical data in order to investigate the strongest interactions between their categories, as well as the directions (trends) of their dispersion. The supplementary variables also used in (MCA) do not participate in the construction of the factorial axes, are simply projected through the transfer relationships together with the other points (categories of variables) in the factorial diagrams and can further contribute to the interpretation of the phenomenon. This paper examines the way in which graphical information can be integrated from the "Interpretive Plane" into the factorial planes that include complementary points. Through the Interpretive Plane, the important interpretive points of the factorial plane are identified, and then those of those that are the "close" to complementary points. Thus, users who wish to use supplmentary variables in their analyses can achieve better and safer interpretative results-conclusions. The above were applied to the investigation of the rewards system, the performance, development and effectiveness of businesses. The business sector variables and the hierarchical position of employees in the enterprise were taken as supplementary.

© 2024, Συγγραφείς
Άδεια CC-BY-SA 4.0