

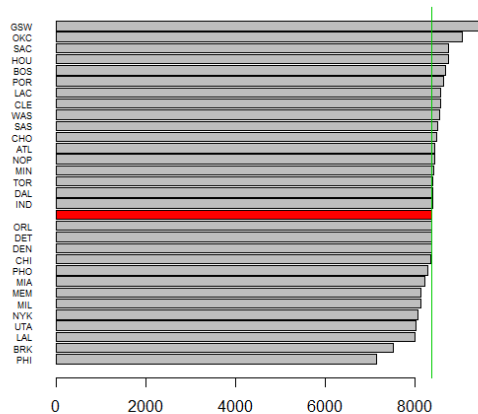
Πρόγραμμα: Ανάλυση Δεδομένων με τη χρήση της R
Επιστημονικός Υπεύθυνος: Καθ. Δημήτριος Καρλής

Τελική Εργασία (Project) Αξιολόγησης

Προκειμένου να αξιολογηθείτε για τις γνώσεις που αποκτήσατε, θα πρέπει να απαντήσετε στα παρακάτω ερωτήματα. Το **παραδοτέο** της εργασίας, το οποίο θα υποβάλετε μέσω της πλατφόρμας ηλεκτρονικής μάθησης, θα πρέπει να περιλαμβάνει τον κώδικα, καθώς και σύντομο σχολιασμό των αποτελεσμάτων.

Τα δεδομένα (τα οποία θα έχετε στη διάθεσή σας σε αρχείο που συνοδεύει την εκφώνηση) αφορούν όλους τους παίκτες του NBA για τη φετινή σεζόν. Με αυτά τα δεδομένα:

1. Διαβάστε τα δεδομένα.
2. Αφαιρέστε γραμμές που έχουν έστω και ένα NA.
3. Για κάθε ομάδα βρείτε το συνολικό αριθμό πόντων.
4. Για κάθε ομάδα βρείτε τον αριθμητικό μέσο πόντων.
5. Για κάθε ομάδα βρείτε το συντελεστή μεταβλητότητας για τους πόντους.
6. Ποιες είναι οι 5 πρώτες ομάδες με βάση τον αριθμό παικτών που χρησιμοποίησαν;
7. Για τα 3,4,5 φτιάξτε κατάλληλα γραφήματα.
8. Υπολογίστε το συντελεστή Gini για τους πόντους και τα λεπτά παιχνιδιού για κάθε ομάδα.
9. Αφαιρέστε τώρα τους παίκτες που έπαιξαν κάτω από 5 λεπτά κατά μέσο όρο. Με αυτούς που απέμειναν υπολογίστε πάλι το συντελεστή Gini για τους πόντους και τα λεπτά παιχνιδιού για κάθε ομάδα. Σχολιάστε.
10. Κάντε ιστογράμματα για τα ποσοστά ευστοχίας στα σουτ 1,2 και 3 πόντων στο ίδιο γράφημα.
11. Κάντε scatterplot για όλους του παίκτες με τα ποσοστά ευστοχίας στα σουτ 2 και 3 πόντων, αφαιρώντας όσους έπαιξαν κάτω από 5 λεπτά ανά παιχνίδι και είχαν τουλάχιστον 20 σουτ 3 πόντων. Για τους 10 παίκτες με το καλύτερο ποσοστό στα σουτ 3 πόντων εμφανίστε το όνομά τους στο γράφημα.
12. Βρείτε τους 5 καλύτερους PG σε σχέση με τον συνολικό αριθμό rebounds που πήραν.
13. Εμφάνισε για αυτούς τους 5 όλα τους τα στατιστικά.
14. Φτιάξτε ένα barplot για τους πόντους κάθε ομάδας. Θα πρέπει να είναι σε φθίνουσα σειρά από πάνω προς τα κάτω και να υπάρχει και ο μέσος όρος πόντων. Όλες οι μπάρες εκτός του μέσου όρου να είναι γκρι ενώ ο μέσος όρος κόκκινος. Δηλαδή το γράφημα να μοιάζει με αυτό.



15. Πόσοι παίκτες σκόραραν πάνω από 1000 πόντους;
16. Πιστεύετε ότι σκοράρουν οι PG τό ίδιο με τους PF; Κάντε τον κατάλληλο έλεγχο.
17. Τρέξτε μια παλινδρόμηση που να συνδέει το ποσοστό επιτυχίας στα σουτ 2 πόντων με αυτό στα σουτ 3 πόντων. Εξηγήστε το μοντέλο κι ελέξτε ις υποθέσεις τους. Εξηγήστε τα ευρήματά σας με λεπτομέρεια.
18. Κατασκευάστε τουλάχιστον 5 γραφήματα σχετικά με τα δεδομένα που πιστεύετε ότι έχουν ενδιαφέρον και σχολιάστε τα.

Τα δεδομένα υπάρχουν στο αρχείο **NBA16-data-for-final-project.csv**.

Οι στήλες αφορούν με τη σειρά που εμφανίζονται στο αρχείο:

- AA
- Player
- Position (Pos)
- Age
- Team (Tm)
- Games (G)
- Games Started (GS)
- Minutes Played (MP)
- Field Goals (FG)
- Field Goals Attempts (FGA)
- 3points made (3P)
- 3points attempt (3PA)
- 2points made (2P)
- 2points attempt (2PA)
- Free Throws (FT)
- Free Throws Achieved (FTA)
- Offensive REbounds (ORB)
- Defensive REbounds (DRB)
- Total Rebounds (TRB)



- Assists (ASS)
- Steals (STL)
- Blocks (BLK)
- Turnovers (TOV)
- Personal Fouls (PF)
- Points (PTS)