UAB
Universitat Autònoma de Barcelona

Dipòsit digital de documents de la UAB

This is the **published version** of the bachelor thesis:

Garrido Rosas, Lucia. *Anàlisi avançada de dades i aprenentatge automàtic aplicats al bàsquet universitari*. Treball de Final de Grau (Universitat Autònoma de Barcelona), 2026 (Grau en Enginyeria de Dades)

This version is available at https://ddd.uab.cat/record/337353

under the terms of the BY-NC-ND license.

Anàlisi avançada de dades i aprenentatge automàtic aplicats al bàsquet europeu

Lucia Garrido Rosas

Resum—Aquest treball analitza el rendiment de jugadors, quintets i equips en situacions de *clutch* a l'Eurolliga, definides com els últims cinc minuts de partit amb una diferència igual o inferior a cinc punts, a partir de dades play-by-play de les temporades 2023 i 2024. Es proposa l'INQ (Impacte Net del Quintet), una mètrica pròpia que captura l'impacte col·lectiu dels quintets incorporant la simetria del joc —el que beneficia un equip perjudica el rival— i comptabilitzant totes les accions del joc. Els resultats mostren que el rendiment global i el rendiment en clutch són dimensions independents als tres nivells analitzats, i que l'INQ és una mètrica vàlida i complementària al PIR oficial de l'Eurolliga. Com a extensió, s'implementa un algorisme genètic per recomanar la composició òptima dels quintets en funció del rival i la situació de joc, identificant oportunitats de millora respecte als quintets habituals. Tots els resultats es presenten mitjançant un dashboard interactiu de Power BI.

Paraules clau— Eurolliga · clutch · INQ · PIR · plus/minus · algorisme genètic · quintets · bàsquet europeu · analítica esportiva · play-by-play

Abstract—This work analyses the performance of players, lineups and teams in *clutch* situations in the EuroLeague, defined as the last five minutes of play with a score difference of five points or less, using play-by-play data from the 2023 and 2024 seasons. A novel metric, the INQ (from the Catalan Impacte Net del Quintet), is proposed to capture collective lineup performance by incorporating the symmetry of the game and accounting for all actions. Results show that global and *clutch* performance are independent dimensions across all three levels of analysis, and that the INQ is a valid metric complementary to the official EuroLeague PIR. A genetic algorithm is implemented to recommend optimal lineup compositions based on the opponent and game situation. All results are presented through an interactive Power BI dashboard.

Index Terms— EuroLeague · clutch · INQ · PIR · plus/minus · genetic algorithm · lineups · European basketball · sports analytics · play-by-play

1 Introducció - Context del treball

La digitalització de l'activitat esportiva ha incrementat significativament el volum i la complexitat de les dades generades en competicions professionals. En el bàsquet, cada partit produeix un conjunt extens d'esdeveniments temporals i variables estadístiques que permeten analitzar el joc amb un alt nivell de detall. Entre aquests escenaris, els anomenats *moments de clutch* [1]— definits com els últims cinc minuts de partit en què la diferència en el marcador és igual o inferior a cinc punts— representen contextos d'alta pressió competitiva i especial interès analític.

Els moments de *clutch* són especialment importants perquè sovint decideixen el resultat final dels partits i poden revelar patrons tàctics i comportament dels jugadors que no es detecten en altres situacions. L'ús d'anàlisi avançada de dades i aprenentatge automàtic permet explorar el rendiment individual, les decisions tàctiques dels equips i l'efecte de factors contextuals, identificant patrons recurrents i suggerint estratègies òptimes

Per dur a terme aquest estudi, s'ha utilitzat un dataset de **play-by-play** corresponent a les temporades 2023 i 2024 de l'Eurolliga. Aquest conjunt de dades registra cada acció durant els partits —com tirs, rebots, assistències, pèrdues de pilota i faltes— amb informació temporal i contextual que permet analitzar amb precisió els moments de *clutch*. La disponibilitat de dades de dues temporades consecutives facilita la comparació de tendències, la validació de models i l'exploració de l'evolució del rendiment dels jugadors i equips en situacions de màxima pressió competitiva.

2 Motivacions

L'esport, i en particular el bàsquet, ha estat un àmbit d'interès rellevant en aquest treball. La pràctica continuada durant diversos anys ha permès adquirir una comprensió de la complexitat tàctica del joc, així com de la importància del treball en equip i de la presa de decisions en moments de màxima pressió. Aquest coneixement, combinat amb l'interès per les matemàtiques i l'enginyeria, ha motivat l'anàlisi del rendiment esportiu des d'una perspectiva quantitativa.

L'aplicació de mètodes analítics i models matemàtics a dades esportives permet integrar aquests dos àmbits. En aquest sentit, el projecte es planteja com una oportunitat per transformar l'experiència pràctica en coneixement més objectiu i estructurat, amb la finalitat de contribuir a una millor comprensió del joc i de les seves dinàmiques. Així mateix, es busca aportar informació útil per a clubs i entrenadors en la presa de decisions sobre la configuració de la plantilla o dels quintets en cada partit.

3 OBJECTIUS

En aquesta secció es veuran els objectius del projecte.

3.1 Objectiu general

Avaluar el rendiment de jugadors i equips en els moments de *clutch* de l'Eurolliga, identificant patrons d'efectivitat i d'oportunitats de millora mitjançant anàlisi de dades i tècniques d'aprenentatge automàtic.

3.2 Objectius específics

1. Determinar els jugadors, quintets i equips amb millor rendiment en clutch mitjançant una mètrica d'eficiència específica.
2. Comparar l'eficiència individual en moments de *clutch* i en la resta del partit amb una mateixa mètrica d'avaluació.
3. Estudiar la relació entre el rendiment en *clutch* i els resultats globals dels partits i de la temporada.
4. Proposar estratègies de millora per equips amb rendiment baix en *clutch* a partir de l'anàlisi de la composició òptima dels quintets mitjançant l'algorisme genètic.
5. Analitzar l'evolució del rendiment i comparació de tendències dels jugadors i equips al llarg de les temporades.
6. Crear visualitzacions i dashboards interactius que facilitin la interpretació de dades i el suport a la presa de decisions d'entrenadors i analistes.

4 METODOLOGIA I PLANTEJAMENT

Aquest treball s'enfoca en l'anàlisi del rendiment dels jugadors i dels equips a la Eurolliga, amb especial atenció als moments de màxima pressió durant els partits. La metodologia s'organitza en diferents tasques que permeten abordar l'estudi de manera ordenada i sistemàtica. Les dades s'obtenen mitjançant l'API pública de l'Eurolliga i s'apliquen tècniques d'anàlisi exploratòria, mètriques d'avaluació pròpies i models d'aprenentatge automàtic. Per oferir una visió general de la planificació, el cronograma del projecte s'ha representat mitjançant un diagrama de Gantt, que es pot consultar a l'Annex A1.

5 ESTAT DE L'ART

La intersecció entre la ciència de dades i el bàsquet professional ha donat lloc a un camp d'investigació actiu que abasta des de l'avaluació del rendiment individual i col·lectiu fins a la predicció de resultats i la presa de decisions tàctiques. A continuació es revisen els principals treballs relacionats amb els cinc eixos temàtics d'aquest projecte.

5.1 Digitalització i analytics en l'esport professional

La digitalització ha transformat profundament la indústria esportiva durant les últimes dècades, modificant no només la manera en què es practiquen els esports, sinó també com s'organitzen, s'analitzen i es consumeixen. Les tecnologies digitals han penetrat en tots els àmbits de l'esport professional, des de la gestió administrativa fins a l'anàlisi del rendiment i la presa de decisions tàctiques. El punt d'inflexió en l'adopció de l'analítica de dades s'atribueix sovint al cas documentat per [3] sobre l'equip de beisbol Oakland Athletics, que va demostrar que les decisions basades en dades podien superar la intuïció i el coneixement tàcit tradicional en la selecció i composició d'equips. Des d'aleshores, l'ús d'eines d'anàlisi s'ha estès progressivament a tots els àmbits de l'esport d'elit, des del reclutament de jugadors fins a la preparació tàctica i la prevenció de lesions.

L'adopció de tecnologies de seguiment i sensors ha fet que informació que abans era inaccessible —com els patrons de moviment en temps real o la càrrega física dels jugadors— es pugui ara recollir i analitzar de manera sistemàtica. Un exemple destacat és la NBA, que va introduir tecnologia de seguiment capaç de registrar els moviments de cada jugador fins a vint-i-cinc vegades per segon, generant estadístiques innovadores sobre el rendiment individual i col·lectiu. De manera similar, durant la Copa del Món de futbol de 2014, la selecció alemanya va utilitzar programari analític per analitzar el rendiment propi i el dels rivals, fet que es considera un factor clau en el seu èxit final [4].

En definitiva, la digitalització ha redefinit la manera en què l'esport professional genera i utilitza el coneixement, desplaçant la presa de decisions des de la intuïció cap a l'evidència quantitativa. Això ha donat lloc a l'aparició de nous perfils professionals dins de les organitzacions esportives, com ara analistes de dades i científics de l'esport, i ha obert un camp d'investigació actiu en la intersecció entre la ciència de dades i el rendiment esportiu [4].

5.2 Models de Machine Learning aplicats al bàsquet

L'aprenentatge automàtic ha esdevingut una eina fonamental en l'anàlisi del rendiment esportiu, ja que permet processar grans volums de dades i descobrir patrons que escapen a l'anàlisi estadística tradicional. En el context del bàsquet, s'ha aplicat a tasques tan diverses com la predicció de resultats de partits, la classificació de jugadors per perfil de joc, la identificació de factors associats al rendiment individual i col·lectiu, o la detecció de situacions crítiques durant els partits. [5]

Els algorismes més habitualment emprats en aquest àmbit inclouen la regressió logística, el Random Forest i les xarxes neuronals, entre d'altres. La regressió logística s'ha utilitzat per classificar jugades i predir resultats gràcies a la seva interpretabilitat; el Random Forest ha demostrat bons resultats en la identificació de factors de rendiment individual per la seva capacitat de gestionar moltes variables simultàniament; i les xarxes neuronals s'han aplicat en tasques més complexes com la detecció de patrons tàctics seqüencials. [6] Els treballs més recents han incorporat mètodes d'interpretabilitat com SHAP [5], que permeten identificar quines variables estadístiques tenen més influència en els resultats dels models, fent les conclusions

més accionables per a entrenadors i analistes.

Aquests avenços han obert la porta a estudis cada vegada més específics, centrats no només en el rendiment global dels jugadors i equips, sinó també en situacions concretes del joc, fent del Machine Learning una eina cada vegada més present en la presa de decisions esportives. [6] Aquesta tendència s'ha estès progressivament al bàsquet europeu, on estudis aplicats a l'Eurolliga han utilitzat models de Machine Learning per analitzar patrons de tir i identificar els factors situacionals que determinen l'èxit d'un llançament a cistella. [7]

5.3 Anàlisi de dades en bàsquet europeu

L'Eurolliga és la competició de clubs de referència a Europa i ha estat objecte d'un interès creixent en la recerca en sports analytics. A diferència de la NBA, que disposa d'una llarga tradició d'estudis analítics, la recerca centrada específicament en el bàsquet europeu és relativament recent, impulsada en gran mesura per la disponibilitat creixent de dades públiques que permeten accedir a informació detallada de cada partit, incloent dades play-by-play, estadístiques de jugadors i resultats.

Des d'una perspectiva comparativa, diversos estudis han analitzat les diferències entre la NBA i l'Eurolliga. Les diferències quantitatives entre ambdues competicions han disminuït de manera consistent al llarg dels anys, essent el ritme de joc la diferència més evident: l'NBA presenta més possessions per partit, mentre que l'Eurolliga es caracteritza per un joc més tàctic i posicional, amb possessions més llargues i major èmfasi en el joc defensiu. [8]

En el context específic de l'Eurolliga, recerques recents han comparat l'efectivitat de diversos indicadors clau de rendiment (KPIs) entre l'Eurolliga i les lligues nacionals europees, identificant que mètriques com el percentatge de tir efectiu, el rating defensiu i la ràtio assistències-pèrdues varien significativament en el seu impacte segons la competició [9], i s'han desenvolupat plataformes específiques per facilitar l'accés i la visualització de dades del bàsquet europeu [10].

5.4 Moments de *clutch* i rendiment sota pressió

Els moments de *clutch*, tal com s'han definit anteriorment, representen els instants de màxima pressió competitiva en un partit de bàsquet i han estat àmpliament estudiats en el camp dels sports analytics. La capacitat de rendir sota pressió és una de les qualitats més valorades en el bàsquet professional, i la seva quantificació ha estat un repte per a investigadors i analistes durant dècades. [1]

Un dels debats centrals en la recerca sobre el *clutch* és si existeixen realment jugadors que rendeixin de manera consistent millor en aquests moments, o si es tracta d'una percepció subjectiva amplificada pels mitjans de comunicació. Estudis basats en dades reals han demostrat que, tot i que alguns jugadors mostren una millora estadísticament significativa del seu rendiment en els moments finals de partits ajustats, la majoria tendeix a mantenir o fins i tot empitjorar lleugerament el seu nivell respecte al global del partit. [11] En particular, s'ha observat que els jugadors solen rendir pitjor en tirs lliures en situacions de màxima pressió, evidenciant un efecte de pressió psicològica mesurable. [12]

Des d'una perspectiva metodològica, s'han proposat diverses mètriques per quantificar el rendiment en *clutch*. Una de les més recents és l'Estimation of Clutch Competency (EoCC), que avalua l'eficiència individual dels jugadors durant els minuts finals de partits ajustats al llarg de múltiples temporades, i ha estat validada comparant els seus resultats amb els vots al premi NBA Clutch Player of the Year. [13] D'altres estudis han utilitzat el True Shooting Percentage (TS%) com a indicador d'eficiència anotadora en situacions de pressió, per la seva capacitat de combinar tirs de camp, triples i tirs lliures en una sola mètrica. [13]

Malgrat l'abundant recerca en el context de la NBA, l'estudi dels moments de *clutch* a l'Eurolliga és pràcticament inexistent en la literatura acadèmica, fet que representa un buit rellevant que aquest treball pretén abordar. Les particularitats del bàsquet europeu, com ara el ritme de joc més tàctic i la major igualtat entre equips, fan que els moments de *clutch* tinguin un paper especialment determinant en el resultat final dels partits. [8]

5.5 Mètriques d'avaluació de quintets

L'avaluació del rendiment col·lectiu dels quintets ha estat un dels reptes més complexos en l'anàlisi del bàsquet professional, ja que requereix capturar no només les contribucions individuals de cada jugador, sinó també les interaccions i sinergies que es produeixen quan cinc jugadors concrets comparteixen pista. Les estadístiques tradicionals de box-score, centrades en accions individuals, resulten insuficients per a aquest propòsit.

La mètrica més bàsica per avaluar l'impacte d'un jugador o quintet és el plus/minus (+/-), que mesura la diferència en el marcador mentre un determinat jugador o conjunt de jugadors és en pista. [14] Tot i la seva simplicitat i intuïtivitat, el plus/minus té limitacions importants: està molt influenciat per la qualitat dels companys i rivals, és sensible al context del partit i requereix mostres grans per ser estadísticament fiable. [14]

Per superar aquestes limitacions, es va desenvolupar l'Adjusted Plus/Minus (APM) [15], que utilitza regressió lineal sobre totes les possessions per estimar la contribució de cada jugador al marcador controlant la qualitat dels companys i rivals en pista. A diferència del plus/minus bàsic, l'APM intenta aïllar l'impacte individual del context del quintet, oferint una estimació més justa de la contribució real de cada jugador. Posteriorment, el Regularized Adjusted Plus/Minus (RAPM) [15] va introduir tècniques de regularització per penalitzar estimacions massa extremes, reduint el soroll estadístic especialment per a jugadors amb pocs minuts. Estudis més recents han estès aquest marc analític per avaluar quintets complets

com a unitat, combinant estadístiques de box-score amb models bayesians per obtenir estimacions més robustes del rendiment col·lectiu. [15]

6 DESENVOLUPAMENT

En aquest apartat es detallen les tasques dutes a terme al llarg del projecte, seguint l'ordre establert en la planificació. El desenvolupament s'estructura en quatre blocs principals: la recollida i preparació de les dades, la modelització i avaluació, l'anàlisi avançada del *clutch* i la visualització dels resultats. Cada bloc descriu les decisions tècniques adoptades, les eines utilitzades i els resultats intermedis obtinguts en cada fase.

6.1 Recollida i preparació de les dades

Les dades de *play-by-play* s'obtenen mitjançant l'API pública de la Eurolliga [2], que permet accedir a informació detallada de cada partit. Aquesta tasca inclou la integració i organització dels datasets, la neteja i unificació de les dades, així com una anàlisi exploratòria inicial. També es defineix la variable associada a les situacions de *clutch* i s'identifiquen aquests moments dins dels partits.

A partir del dataset original de play-by-play, s'han dut a terme diverses tasques de neteja, correcció i enriquiment de les dades per garantir la qualitat i la consistència del conjunt final utilitzat en l'anàlisi. Una visió esquemàtica del flux complet de processament de dades es pot consultar a l'Annex A2.

Pel que fa a la correcció d'errors, s' han normalitzat jugades amb valors incorrectes al camp MARKERTIME, completat les jugades d'inici de quart sense temps registrat i unificat els noms de jugadors amb grafies diferents. Quant a l'enriquiment, s'han afegit un identificador únic de partit (GAMEID); identificadors numèrics de tipus de jugada (PLAYID, PLAYEXTRA); els codis dels equips local i visitant (Team_A, Team_B); una variable booleana per a les situacions de *clutch* (CLUTCH); els valors acumulats de l'INQ per equip (INQ_total_A/B, INQ_clutch_A/B); i les columnes de quintets en pista (Lineup_A, Lineup_B) i l'identificador de quintet (lineup_id), obtinguts d'un dataset complementari de la mateixa API. El model relacional complet es pot consultar a l'Annex A3.

6.1.1. Mètriques pròpies: INQ i INQ Individual

En el marc d'aquest projecte s'han definit dues mètriques pròpies per avaluar el rendiment en situacions de *clutch*: l'INQ (Impacte Net del Quintet), orientada a mesurar l'impacte col·lectiu dels quintets, i l'INQ-individual, que estén aquesta mateixa lògica a nivell de jugador.

Les mètriques existents per avaluar el rendiment en bàsquet presenten limitacions importants quan s'apliquen a l'anàlisi col·lectiva. D'una banda, la valoració estàndard europea —el PIR (Performance Index Rating)— mesura el rendiment individual de cada jugador sumant les seves accions positives i restant les negatives, però no està dissenyada per avaluar quintets com a unitat ni captura l'efecte que les accions d'un equip tenen sobre el rival. D'altra banda, el plus/minus (+/−) aproxima l'impacte col·lectiu però ignora totes les accions que no es tradueixen directament en punts —rebots, robatoris, taps, assistències o pèrdues— i és molt sensible al context del partit.

L'INQ neix precisament per cobrir aquest buit: és una variant del PIR adaptada per avaluar quintets en lloc de jugadors individuals, que incorpora la simetria del +/− —el que beneficia un equip perjudica l'altre— però comptabilitzant totes les accions del joc amb els pesos de la valoració estàndard, no només els punts. Això permet capturar contribucions col·lectives que ni el PIR ni el +/− reflecteixen adequadament, com la pressió defensiva que força una pèrdua rival o el bloqueig que allibera un company per agafar un rebot. Ambdues mètriques es calculen tant per al global del partit com per exclusivament als moments de *clutch*.

INQ: Impacte Net del Quintet

L'INQ assigna a cada acció del joc un valor numèric que afecta simultàniament al quintet propi i al rival, seguint la lògica de la valoració estàndard europea adaptada a una perspectiva col·lectiva i simètrica. Els tirs anotats reben un valor proporcional als punts generats; les accions que mantenen o recuperen la possessió (rebots, robatoris i taps) valen +1 per al quintet propi i −1 per al rival; i les que la cedeixen (tirs fallats, pèrdues i faltes) apliquen la penalització inversa. Les assistències s'assignen de manera simètrica (+1/−1) perquè premien el joc col·lectiu i penalitzen la incapacitat defensiva rival d'evitar un tir en condicions òptimes. La falta rebuda i el tap rebut queden implícits en les accions del rival, evitant el doble comptatge. La taula següent recull l'impacte de cada acció:

Acció	INQ propi	INQ rival
Tir de 2 anotat	2	−2
Tir de 3 anotat	3	−3
Tir lliure anotat	1	−1
Tir fallat	−1	1
Rebot ofensiu	1	−1
Rebot defensiu	1	−1
Robatori	1	−1
Tap	1	−1
Assistència	1	-1
Pèrdua	−1	1
Falta comesa	−1	1

Cal tenir en compte que l'INQ és una mètrica de rendiment que no controla la qualitat dels rivals ni el context del partit. S'ha d'interpretar sempre de manera comparativa i en el context de les situacions analitzades.

INQ Individual

L'INQ individual es calcula com la mitjana ponderada per possessions de l'INQ dels quintets en què ha participat el jugador, al llarg de tots els partits de la temporada. Aquesta mètrica permet identificar jugadors que aporten valor col·lectiu més enllà del que reflecteixen les seves

estadístiques individuals — jugadors que, simplement per estar en pista, fan que el seu quintet funcioni millor.

La ponderació per possessions és necessària perquè un jugador pot haver participat en quintets diferents al llarg del partit, cadascun amb un nombre de possessions distint, i una mitjana simple tractaria igual un quintet que ha jugat 2 possessions que un que n'ha jugat 40, distorsionant el valor final. El càlcul d'aquesta mètrica ve donat per la formula següent:

$$INQ_{\text{individual}} = \frac{\sum_{q \in Q} INQ_q \times poss_q}{\sum_{q \in Q} poss_q}$$

On:

- $q \in Q \rightarrow$ quintets en què ha participat el jugador durant el partit
- $INQ_q \rightarrow$ INQ acumulat del quintet q durant el partit
- $poss_q \rightarrow$ possessions jugades en el quintet q

*L'***INQ-individual clutch** s'obté aplicant exactament la mateixa fórmula però restringint el càlcul exclusivament a les possessions en situació de *clutch*. Això permet comparar directament el rendiment col·lectiu d'un jugador durant tot el partit amb el seu rendiment en els moments de màxima pressió, identificant jugadors que milloren o empitjoren la seva contribució col·lectiva en situacions decisives.

Cal tenir en compte també que l'INQ individual no aïlla la contribució pròpia del jugador de la del seu quintet. Per aquest motiu, s'ha d'interpretar sempre en context i de manera complementària a les estadístiques individuals.

6.1.2. Datasets derivats

A partir del dataset principal, s'han generat dos datasets derivats per facilitar l'anàlisi posterior:

- **player_stats_by_game.csv:** conté estadístiques individuals de jugadors per partit, tant en el global com en situació de *clutch*. Inclou les estadístiques estàndard de box-score (punts, rebots, assistències, pèrdues, robatoris, taps i faltes), el PIR, el plus/minus i els valors d'INQ-individual calculats seguint la lògica descrita a l'apartat 6.1.1. Cada fila correspon a un jugador en un partit concret.
- **lineup_inq.csv:** conté el rendiment de cada quintet desagregat per partit. Inclou l'INQ-global i INQ-clutch acumulats, el nombre de jugades disputades en cada context, el plus/minus del quintet i una variable booleana que indica si el quintet va disputar jugades en situació de *clutch*. Cada fila correspon a un quintet en un partit concret.

A més dels datasets derivats, s'han generat cinc diccionaris de referència en format CSV per facilitar la consulta i el creuament d'informació al llarg de l'anàlisi:

- **teams.csv**: relaciona les sigles de cada equip amb el seu nom complet i el seu logo.
- **players.csv**: conté la llista de jugadors amb el seu equip i temporada corresponent, permetent traçar la trajectòria de cada jugador al llarg de les dues temporades analitzades.
- **gamecode_season.csv**: relaciona cada identificador de partit (GAMEID) amb la temporada, la fase i la ronda corresponents, així com els equips participants.
- **lineups.csv**: assigna un identificador únic (quintet_id) a cada quintet únic que ha aparegut en pista al llarg de les dues temporades, amb els cinc jugadors identificats individualment.
- **play.csv**: mapeja cada descripció textual de jugada del camp PLAYINFO al seu identificador numèric (PLAYID), permetent treballar amb els tipus de jugada de manera estructurada i eficient al llarg de tota l'anàlisi.

El model relacional que il·lustra les relacions entre el dataset principal, els datasets derivats i els diccionaris de referència es pot consultar a l'Annex A3.

6.2 Optimització de quintets mitjançant algorisme genètic

Es proposa l'aplicació d'un algorisme genètic per identificar la composició òptima dels quintets en funció del rival i la situació de joc. Amb plantilles d'entre catorze i vint jugadors i una mitjana de setze jugadors per equip, el nombre de combinacions possibles de cinc jugadors oscil·la entre 2.002 i 15.504, amb una mitjana aproximada de 4.368 combinacions per equip. La majoria d'aquestes combinacions mai han coincidit a pista, de manera que no existeixen dades reals del seu rendiment conjunt. L'algorisme genètic permet explorar aquest espai de manera sistemàtica, estimant el rendiment de quintets no observats a partir de les contribucions individuals dels jugadors i identificant combinacions potencialment més efectives que les utilitzades realment [16].

6.2.1. Fonaments de l'algorisme genètic

Un algorisme genètic és una tècnica d'optimització inspirada en el procés d'evolució natural [17]. A partir d'un conjunt inicial de solucions candidates —la població—, l'algorisme aplica iterativament tres operacions: la selecció, que escull les millors solucions; el creuament, que combina parts de dues solucions per generar-ne de noves; i la mutació, que introdueix canvis aleatoris per afavorir la diversitat. Cada solució s'avalua mitjançant una funció de fitness i la població evoluciona progressivament cap a solucions de qualitat creixent.

Aquesta tècnica és especialment adequada per al problema plantejat perquè opera sobre representacions discretes —com una llista de cinc jugadors— sense imposar cap restricció sobre la forma de la funció objectiu, la qual cosa la fa idònia per a problemes combinatoris com la selecció de quintets [18].

6.2.2. Adaptació a la recomanació de quintets

En el context d'aquest projecte, cada individu de la població representa un quintet de cinc jugadors del roster d'un equip per a un enfrontament concret. La funció de fitness es defineix a partir de les mètriques individuals d'INQ

de cada jugador condicionades al rival, combinant el rendiment estimat davant aquell rival i la seva consistència al llarg de la temporada:

$$f(q, rival) = \sum_{i=1}^{5} [INQ_ponderat(i, rival) - \lambda \cdot INQ_std(i)]$$

On **INQ_ponderat(i, rival)** és una estimació del rendiment del jugador i davant el rival concret, obtinguda mitjançant una ponderació progressiva entre les dades específiques vs rival i el valor global:

$$INQ_ponderat(i, rival) = \alpha(n) \cdot INQ_mean_vs_rival(i) + (1 - \alpha(n)) \cdot INQ_mean_global(i)$$

On:

- $\alpha(n) = \min(n/2, 1)$, amb n= partits disputats pel jugador i contra aquell rival.
- $INQ_mean_vs_rival(i)\rightarrow$ mitjana ponderada per jugades de l'INQ del jugador $i$contra aquell rival
- $INQ_mean_global(i)\rightarrow$ mitjana ponderada per jugades de l'INQ del jugador $i$al llarg de la temporada.
- $INQ_std(i)\rightarrow$ desviació estàndard global del jugador i, mesura de la seva inconsistència
- $\lambda = 0{,}5\rightarrow$ paràmetre de penalització que assigna el mateix pes al rendiment mitjà estimat i a la penalització per inconsistència, reflectint la decisió de considerar la variabilitat del jugador tan rellevant com el seu valor mitjà a l'hora de seleccionar un quintet òptim.

El coeficient α creix progressivament amb el nombre de partits disputats contra el rival: amb zero partits s'utilitza íntegrament el valor global; amb un partit α = 0,50; i a partir de dos partits s'utilitzen exclusivament les dades específiques. El llindar N = 2 respon al fet que la fase regular garanteix exactament dos enfrontaments per parella i temporada, de manera que la gran majoria de casos disposen de dades suficients per usar les dades contra el rival directament. En fase de playoffs, amb fins a set partits possibles, les dades específiques contra el rival acumulen una mostra molt més àmplia, cosa que reforça la fiabilitat de l'estimació del rendiment davant aquell rival concret

D'aquesta manera, la funció de fitness afavoreix jugadors amb un rendiment elevat i estable davant el rival concret. Totes les mètriques es calculen per separat per a dues situacions de joc —clutch i global— a partir del dataset player_stats_by_game.csv, permetent generar recomanacions específiques per a cada context.

6.2.3. Desenvolupament de l'algorisme

Totes les mètriques d'INQ utilitzades per l'algorisme es calculen seguint la mateixa lògica de ponderació per jugades descrita a l'apartat 6.1.1. El roster elegible per a cada combinació equip-temporada-situació es determina aplicant filtres de representativitat: un mínim de 24 partits per a la situació global i un mínim de 8 partits per a la situació *clutch*, determinats a partir d'una anàlisi de representativitat estadística detallada a l'apartat 7.1, garantint que tots els jugadors considerats disposen d'una mostra estadísticament fiable.

La població inicial es construeix a partir dels quintets reals registrats al dataset lineups.csv, complementats amb combinacions generades aleatòriament fins a assolir una mida de 80 individus, garantint la diversitat inicial. Pel que fa als operadors genètics, la selecció es realitza mitjançant torneig, escollint en cada iteració el millor individu d'un subconjunt aleatori de k = 5 individus; el creuament combina jugadors de dos quintets amb una probabilitat de 0,85; i la mutació substitueix un jugador per un altre del mateix roster amb una probabilitat de 0,20.

En cada generació es preserven els tres millors individus per elitisme, garantint que la qualitat de la solució no decreix entre generacions. En tots els casos s'apliquen dues restriccions per garantir que els individus generats siguin vàlids: que no hi hagi jugadors repetits dins del quintet, i que cada jugador pertanyi al roster elegible.

L'algorisme s'executa de manera independent per a cada combinació d'equip, rival, temporada i situació de joc, de manera que les recomanacions reflecteixen tant la plantilla disponible en cada temporada com el context específic de l'enfrontament. El procés s'itera durant un màxim de 150 generacions, tot i que en la pràctica l'algorisme convergeix al voltant de les 100 generacions amb les dades disponibles; les generacions addicionals actuen com a marge de seguretat per garantir l'estabilitat de la solució. Al final del procés s'extreuen els cinc quintets únics amb major fitness com a resultat.

6.2.1. Output de l'algorisme

Per a cada combinació d'equip, rival, temporada i situació de joc, l'algorisme retorna tres elements principals:

El **top-5 de quintets òptims proposats**: les cinc combinacions de jugadors que maximitzen la funció de fitness per a aquell enfrontament concret, amb la seva puntuació estimada i el coeficient α de cada jugador, que indica en quina mesura la recomanació es basa en dades específiques vs rival o en el valor global.

El **quintet real de referència**: el quintet més utilitzat per aquell equip en aquella situació contra aquell rival, identificat com el que acumula el major nombre de jugades. El seu rendiment s'avalua de dues maneres complementàries: aplicant la mateixa funció de fitness que els quintets proposats —garantint una comparació homogènia— i mitjançant l'INQ real registrat, calculat com la mitjana ponderada per jugades de l'INQ del quintet en els partits disputats contra aquell rival.

Els **dos deltes de comparació**:

- δ_fitness = fitness proposat − fitness estimat del quintet de referència → comparació homogènia.
- δ_real = fitness proposat − INQ real registrat del quintet de referència → comparació contextual

Un valor positiu de δ indica que l'algorisme identifica una combinació potencialment millor que la utilitzada realment; un valor negatiu indica que el quintet habitual ja és difícilment superable segons les mètriques disponibles. Aquest output connecta directament amb l'objectiu específic 4 del treball. Els resultats s'exporten al dataset *optimal_lineups_by_rival.csv i* es presenten de manera interactiva al dashboard de Power BI (Annex A6). El pipeline complet de l'algorisme es pot consultar a l'Annex A4.

7 ANÀLISI DE RESULTATS

Aquest apartat presenta els resultats obtinguts al llarg del projecte, estructurats per donar resposta als objectius específics definits a l'apartat 3.2. Els resultats es complementen amb visualitzacions interactives al dashboard de Power BI, que permeten explorar-los amb major detall filtrant per equip, temporada i situació de joc.

7.1 Rànquings de rendiment: jugadors, quintets i equips

Aquesta secció presenta els rànquings de rendiment basats en la mètrica INQ per a tres nivells d'agregació: jugadors individuals (7.1.1), quintets (7.1.2) i equips (7.1.3). Per a cadascun s'analitza tant el rendiment global al llarg de la temporada com el rendiment específic en situacions de *clutch*, permetent identificar quins jugadors, combinacions i equips mantenen un nivell d'efectivitat col·lectiva superior en els moments decisius del partit. Per garantir la fiabilitat estadística dels resultats, s'han aplicat filtres de representativitat diferenciats per a jugadors i quintets, determinats a partir d'una anàlisi exploratòria documentada al notebook *representativity_filters.ipynb*; als equips no s'hi apliquen filtres. Els criteris concrets s'expliquen dins de cada subapartat.

7.1.1. Rànquing de jugadors

Per a cada jugador s'ha calculat l'INQ-individual al llarg de la temporada mitjançant una mitjana ponderada per possessions sobre tots els partits disputats, tal com es descriu a l'apartat 6.1.1:

$$INQ_{\text{jugador}} = \frac{\sum_{p \in P} \sum_{q \in Q_p} INQ_{q,p} \cdot poss_{q,p}}{\sum_{p \in P} \sum_{q \in Q_p} poss_{q,p}}$$

On:

- $p \in P \rightarrow$ partits en què ha jugat el jugador j
- $q \in Q_p \rightarrow$ quintets en què ha participat el jugador $j$durant el partit p
- $INQ_{q,p} \rightarrow$ INQ acumulat del quintet $q$en el partit p
- $pos_{q,p} \rightarrow$ possessions jugades pel quintet $q$en el partit p

L'INQ-clutch s'obté aplicant la mateixa fórmula filtrant exclusivament les possessions en situació de *clutch*.

Per als jugadors, l'anàlisi revela que la desviació estàndard de l'INQ individual no varia significativament en funció del nombre de jugades per partit, ja que la mitjana ponderada per possessions s'estabilitza de manera natural. Per aquest motiu, s'ha optat pel percentil 40 de la distribució de partits jugats com a criteri objectiu i consistent entre els dos contextos, garantint que els jugadors del rànquing han tingut una participació regular al llarg de la temporada: 24 partits en el global i 8 partits *clutch* en el *clutch*, retenant proporcions similars — 61,9% i 60,2% respectivament (N = 361 i N = 303). L'evidència estadística d'aquesta decisió es pot consultar a l'Annex A5.

Els resultats es mostren acompanyats del nombre de partits disputats com a factor de representativitat, i es poden explorar al dashboard de Power BI amb filtres per temporada i equip (Annex A6).

Els rànquings de jugadors revelen una divergència clara entre el rendiment global i el rendiment en *clutch*, confirmant que la capacitat de generar impacte col·lectiu en moments decisius no es correspon necessàriament amb el rendiment general. En el rànquing global, Walter Tavares encapçala la classificació en ambdues temporades — amb un INQ de 9,57 el 2023 (34 partits) i 7,00 el 2024 (40 partits) — consolidant-se com el jugador amb major impacte col·lectiu de la competició. Entre temporades s'observen canvis notables al top: el 2023 destaquen Musa (8,87), Causeur (8,50) i Yabusele (8,16), mentre que el 2024 Mirotic (6,20), Larkin (6,16) i Shorts (5,91) hi entren per primera vegada, reflectint els moviments de plantilla de la competició. Campazzo i Deck mantenen posicions similars en ambdues temporades.

El rànquing *clutch* presenta un perfil substancialment diferent. Larentzakis lidera el 2023 amb un INQ-clutch de 9,49 (15 partits), seguit de Tavares (8,50) i Musa (7,72). El 2024, en canvi, el lideratge passa a Williams (8,04), Grant (7,28) i Hayes-Davis (6,83), mostrant una major rotació al rànquing *clutch* entre temporades. La resta del rànquing *clutch* no coincideix amb el rànquing global en cap de les dues temporades, a excepció de Tavares i Musa el 2023, i Deck el 2024, que apareixen en ambdós rànquings però en posicions inferiors al *clutch*. Això confirma que l'impacte col·lectiu en situacions de pressió és una qualitat diferenciada que les mètriques globals no capturen adequadament.

7.1.2. Rànquing de quintets

Les mètriques principals s'han calculat a partir del dataset **lineup_inq.csv**, tal com s'ha descrit a l'apartat 6.1.2. Per obtenir el rendiment d'un quintet al llarg de la temporada, s'utilitza una mitjana ponderada per possessions sobre tots els partits disputats, seguint la mateixa lògica que per als jugadors: ponderar per possessions garanteix que els períodes amb més activitat tinguin més pes, i calcular la versió *clutch* simplement restringeix el càlcul a les possessions en situació de pressió. Aquesta coherència entre mètriques fa que els rànquings de jugadors, quintets i equips siguin directament comparables entre si.

$$INQ_{quintet} = \frac{\sum_{p \in P_q} INQ_{q,p} \cdot poss_{q,p}}{\sum_{p \in P_q} poss_{q,p}}$$

On:

- $p \in P \rightarrow$ partits en què ha jugat el quintet q
- $INQ_{q,p} \rightarrow$ INQ acumulat del quintet $q$en el partit p
- $poss_p \rightarrow$ possessions jugades en el partit p

Això garanteix que els partits on el quintet ha jugat més tinguin més pes en el resultat final. L'INQ-clutch mitjà aplica la mateixa fórmula però restringint el càlcul als partits en què el quintet ha jugat almenys una jugada en situació de *clutch* i ponderant per les possessions *clutch* jugades.

Per als quintets, s'han establert dos filtres independents per a cada context. Per al global, s'ha aplicat un mínim de 2 partits disputats (~P70), ja que la distribució de partits és molt concentrada en valors baixos i aquest llindar garanteix que el rendiment no correspon a un únic encontre puntual; aquest filtre reté el 46,6% dels quintets (N = 5.593). Per al *clutch*, s'ha fixat el mateix mínim de 2 partits *clutch* com a criteri metodològic, donat que la distribució *clutch* és molt concentrada en 1 partit; aquest filtre reté el 22,2% dels quintets (N = 504). Els gràfics que sustenten aquesta decisió es troben a l'Annex A5. A més dels rànquings, el dashboard de Power BI permet explorar l'evolució de l'INQ dels quintets al llarg de les possessions d'un partit concret, filtrant per quarts i situació, així com consultar en detall les estadístiques individuals dels jugadors del quintet seleccionat (Annex A6).

Els rànquings de quintets mostren, de manera similar als jugadors, una divergència notable entre el rendiment global i el rendiment en *clutch*, amb combinacions de jugadors clarament diferenciades en cada context.

En el rànquing global, el quintet líder de 2023 és De Colo, Jackson, Lauvergne, Lee i Scott (ID 10486, INQ = 37,04), mentre que el 2024 ho és Cordinier, Grazulis, Morgan, Pajola i Polonara (ID 9566, INQ = 38,73). El segon quintet del 2024, Grant, Hernangómez, Kalaitzakis, Lessort i Nunn (ID 13064, INQ = 37,49), és especialment destacable per la seva proximitat al líder.

En el rànquing *clutch*, el quintet líder del 2023 és Canaan, Milutinov, Peters, Walkup i Williams-Goss (ID 8591, INQ-clutch = 24,00), seguit d'Abrines, Kalinic, Laprovittola, Satoransky i Vesely (ID 698, INQ-clutch = 23,15). El 2024, el lideratge clutch correspon a De Colo, Lighty, Maledon, Roberson i Sako (ID 10593, INQ-clutch = 18,91), seguit de Grant, Hernangómez, Kalaitzakis, Lessort i Sloukas (ID 13066, INQ-clutch = 18,65) — una variant del segon quintet global amb Sloukas en lloc de Nunn.

Cap dels quintets líders en *clutch* coincideix amb els líders globals, confirmant que les combinacions més efectives en moments de pressió responen a un perfil diferent del que maximitza el rendiment al llarg del partit. En conjunt, els resultats confirmen que el rendiment global i el rendiment *clutch* dels quintets són dimensions independents que requereixen anàlisi diferenciada.

7.1.3. Rànquing d'equips

Per determinar quin equip ha rendit millor al llarg de la competició, tant en termes generals com en situacions de *clutch*, s'ha calculat l'INQ d'equip. Aquesta mètrica s'obté aplicant la mateixa fórmula de mitjana ponderada per possessions utilitzada per als jugadors i els quintets, però en lloc de filtrar per un jugador o quintet concret, s'agafen totes les files del dataset **lineup_inq.csv** corresponents a l'equip — és a dir, tots els seus quintets en tots els partits de la temporada:

$$INQ_{equip} = \frac{\sum_{q \in Q} \sum_{p \in P_q} INQ_{q,p} \cdot poss_{q,p}}{\sum_{q \in Q} \sum_{p \in P_q} poss_{q,p}}$$

On:

- $q \in Q \rightarrow$ quintet de l'equip
- $p \in P_q \rightarrow$ partits en què ha jugat el quintet q
- $INQ_{q,p} \rightarrow$ INQ acumulat del quintet $q$en el partit p
- $poss_{q,p} \rightarrow$ possessions jugades pel quintet $q$en el partit p

L'INQ-clutch d'equip s'obté aplicant la mateixa fórmula filtrant exclusivament les possessions en situació de *clutch*. A diferència dels jugadors i els quintets, als equips no s'apliquen filtres de representativitat. La raó és que mentre un jugador concret pot haver disputat pocs partits o un quintet específic pot haver coincidit en pista en molt poques ocasions, cada equip acumula necessàriament les dades de tots els seus quintets al llarg de tota la temporada, de manera que la seva mostra és sempre completa i comparable entre equips. A més, la ponderació per possessions ja s'encarrega de minimitzar l'impacte dels quintets que han jugat poc, fent que el seu pes en el valor final de l'equip sigui proporcionalment petit i no distorsioni el resultat. Els resultats es mostren al dashboard de Power BI on és possible filtrar per temporada (Annex A6).

El rànquing d'equips revela que el Real Madrid lidera l'INQ-global en ambdues temporades, tot i que amb una caiguda notable entre el 2023 (5,84) i el 2024 (3,82), suggerint una competició més igualada en la segona temporada. El Panathinaikos es manté en segona posició al rànquing basat en l' INQ en les dues temporades tant en global com en *clutch*, consolidant-se com el segon equip més consistent de la competició.

La divergència entre rendiment global i *clutch* és especialment pronunciada en alguns equips. L'EA7 Milano presenta un INQ-global positiu en ambdues temporades però negatiu en *clutch* el 2023 (−2,63), indicant que el seu rendiment es degrada significativament en els moments decisius. En sentit contrari, l'Olympiacos el 2023 presenta un INQ-global de 2,60 però un INQ-clutch de 3,43, suggerint un equip especialment competitiu en moments de pressió.

L'ALBA Berlin és l'únic equip amb INQ-global negatiu

en les dues temporades (−1,16 el 2023 i −1,98 el 2024), mentre que el 2024 passa a tenir INQ-clutch positiu (0,54), una millora notable respecte al −1,35 del 2023.

La mitjana global de la competició cau de 2,01 el 2023 a 1,65 el 2024, mentre que la de *clutch* es manté estable (1,41 el 2023 i 1,37 el 2024), indicant que la competició s'ha igualat en joc general però els moments decisius mantenen un nivell consistent.

En conjunt, els resultats confirmen que, d'igual manera que en els jugadors i els quintets, el rendiment global i el rendiment en *clutch* dels equips són dimensions independents

7.2. Validació de l'INQ: capacitat predictiva i relació amb el PIR.

Per tal de validar l'INQ com a mètrica pròpia, es compara el seu poder predictiu amb el del PIR (Performance Index Rating), la mètrica oficial de l'Eurolliga, tant sobre el resultat dels partits com sobre la classificació final de la fase regular. L'anàlisi completa, incloent els càlculs, les visualitzacions i les taules de resultats, es pot consultar al notebook **analisi_inq_classificacio.ipynb** (Annex A7).

Capacitat predictiva sobre el resultat dels partits

S'ha calculat, per a cada mètrica, el percentatge de partits en què l'equip amb el valor superior guanya. Els resultats mostren que el PIR és la mètrica que millor discrimina el resultat (90,2%), seguida de l'INQ-global (81,5%), el PIR-clutch (79,9%) i l'INQ-clutch (75,1%), tots molt per sobre del 50% esperat per atzar.

Per complementar aquesta anàlisi, s'ha calculat la correlació puntibiserial entre cada mètrica i el resultat del partit — una mesura estadística adequada quan es relaciona una variable contínua amb una variable binària com el resultat (victòria o derrota). Els valors obtinguts mantenen el mateix ordre que el percentatge de partits guanyats, amb l'única excepció de l'intercanvi entre INQ-clutch i PIR-clutch, amb valors molt propers (0,531 vs 0,507): PIR ($r = 0,617$), INQ-global ($r = 0,577$), INQ-clutch ($r = 0,531$) i PIR-clutch ($r = 0,507$). En qualsevol cas, totes quatre mètriques presenten correlacions estadísticament significatives ($p < 0,001$).

Correlació amb la classificació final i error de predicció

Per analitzar si les mètriques reflecteixen el rendiment a llarg termini, s'ha calculat la correlació entre la mitjana de cada mètrica al llarg de la fase regular i la posició final a la classificació. El PIR presenta la correlació més alta, seguit de l'INQ-global, l'INQ-clutch i el PIR-clutch ($r = −0,485$, $p < 0,001$ per a tots). Per quantificar aquesta capacitat predictiva en termes més interpretables, s'ha calculat l'error absolut mitjà (MAE) entre el rànquing de cada mètrica i la classificació real. El PIR obté un MAE de 3,89 posicions el 2023 i 3,06 el 2024, mentre que l'INQ-global obté 4,00 i 3,82 respectivament — una diferència mínima que s'estreny considerablement en la segona temporada. Ambdues mètriques milloren la seva correlació el 2024 (INQ-global: $r = −0,662$; PIR: $r = −0,727$) respecte al 2023 (INQ-global: $r = −0,525$; PIR: $r = −0,557$).

L'anàlisi de correlació amb la classificació final s'ha realitzat únicament amb l'INQ-global, ja que no tots els partits de la fase regular contenen situacions de clutch, fet que redueix la mostra disponible per a l'INQ-clutch i en limita la representativitat per reflectir el rendiment global de la temporada

Els casos més destacats il·lustren tant la capacitat com els límits de la mètrica. L'Olympiacos el 2023, líder de la classificació amb un INQ-global entre els més alts de la competició, és un exemple on la mètrica reflecteix fidelment el rendiment col·lectiu i el resultat final. En canvi, el Panathinaikos el 2023 — amb l'INQ-global més alt de la competició (~5,5) però classificat 17è — representa un cas on un rendiment col·lectiu elevat no es tradueix en resultats competitius, possiblement per factors contextuals com lesions, canvis de plantilla o situacions puntuals que les mètriques de rendiment brut no capturen.

A més, el dashboard de Power BI permet visualitzar l'evolució de l'INQ al llarg de les possessions d'un partit concret per als dos equips, amb la possibilitat de filtrar per quart i per situació de *clutch*, permetent relacionar el flux de rendiment col·lectiu amb el resultat final del partit (Annex A6).

Relació entre mètriques

La matriu de correlació entre les quatre mètriques revela que l'INQ-global i el PIR presenten una correlació moderada ($r = 0,366$), indicant que, tot i mesurar el rendiment en el mateix context, capturen dimensions diferenciades: el PIR agrega accions individuals mentre que l'INQ mesura l'impacte col·lectiu del quintet com a unitat. Aquest resultat justifica l'INQ com a mètrica complementària al PIR i no com a substitutiva. En canvi, l'INQ-clutch i el PIR-clutch mostren una correlació substancialment més alta ($r = 0,694$), suggerint que ambdues convergeixen en la captura del rendiment en situacions de pressió.
En conjunt, aquests resultats validen l'INQ com a mètrica complementària al PIR: amb una capacitat predictiva del 81,5% sobre el resultat dels partits i una correlació moderada amb el PIR ($r = 0,366$), l'INQ aporta informació diferenciada sobre el rendiment col·lectiu que les mètriques individuals no capturen.

7.3. Recomanació de quintets

Els resultats de l'algorisme genètic es presenten al dashboard de Power BI, on l'usuari pot seleccionar qualsevol combinació d'equip, rival, temporada i situació de joc. A continuació s'analitzen tres tipus de resultats representatius.

Casos amb millora potencial significativa

Un exemple representatiu és el Baskonia contra el Virtus Bologna en la temporada 2023 en situació de *clutch*. L'AG proposa el quintet format per Moneke, Kotsar,

Marinkovic, Miller-McIntyre i Sedekerskis (fitness = −6,88), mentre que el quintet de referència —Costello, Howard, Marinkovic, Miller-McIntyre i Moneke— obté un fitness estimat de −16,82 (δ_fitness = +9,94 · δ_real = +5,12). Tot i que ambdós quintets presenten fitness negatiu, l'AG identifica una combinació significativament menys perjudicial per a aquest enfrontament concret.

Casos on el quintet habitual és ja òptim

L'Istanbul contra el Zalgiris en la temporada 2024 en situació de *clutch* presenta un δ_fitness de 0,00: l'AG proposa exactament el mateix quintet que l'equip ja utilitza —Bryant, Dozier Jr, Larkin, Oturu i Smits— amb un fitness de 12,61, confirmant que la decisió tàctica de l'entrenador és òptima segons les mètriques disponibles.

Casos on el quintet habitual supera la recomanació

El Milan contra el Zalgiris en la temporada 2024 en situació de *clutch* obté un δ_fitness de −0,24, indicant que el quintet habitual —Bolmaro, Leday, Mannion, Shields i Tonut— supera lleugerament el millor quintet proposat per l'AG (fitness = 5,06 vs 5,31). Aquest resultat posa de manifest que les sinergies entre jugadors que coincideixen habitualment en pista no sempre queden recollides en la funció de fitness, reforçant la utilitat de l'AG com a eina de suport i no com a substitut del criteri de l'entrenador.

8 Conclusions

Aquest treball ha abordat l'anàlisi avançada del rendiment en situacions de *clutch* a l'Eurolliga, combinant el disseny de mètriques pròpies, l'anàlisi estadística i l'aplicació d'un algorisme genètic per a la recomanació de quintets. Els resultats obtinguts permeten extreure les conclusions següents.

L'INQ ha demostrat ser una mètrica vàlida i complementària al PIR oficial, amb una capacitat de discriminar el resultat dels partits del 81,5% en la versió global i del 75,1% en la versió *clutch*. La correlació moderada entre ambdues (r = 0,366) confirma que mesuren dimensions diferenciades, justificant l'INQ com a eina complementària i no substitutiva.

L'anàlisi ha confirmat que el rendiment global i el rendiment en *clutch* són dimensions independents als tres nivells analitzats, i que la capacitat de rendir en moments decisius no es reflecteix necessàriament en les mètriques generals. Casos com el Panathinaikos el 2023 —2n per INQ però 17è a la classificació— evidencien que un bon rendiment col·lectiu no sempre es tradueix en resultats competitius, possiblement per factors contextuals que les mètriques de rendiment brut no capturen.

L'algorisme genètic ha identificat oportunitats de millora reals en determinats enfrontaments, alhora que confirma l'optimalitat del quintet habitual quan escau. La diversitat de resultats reflecteix que l'efectivitat d'un quintet depèn també de sinergies col·lectives que la funció de fitness no captura completament, reforçant la seva utilitat com a eina de suport i no com a substitut del criteri de l'entrenador.

Pel que fa a les limitacions, l'INQ és una mètrica de rendiment brut que no controla la qualitat dels rivals ni les sinergies entre jugadors, i la funció de fitness de l'AG és additiva per naturalesa, de manera que la suma de les contribucions individuals pot subestimar o sobreestimar el rendiment real d'un quintet com a unitat. A més, restringir l'anàlisi a dues temporades —necessari perquè els rosters canvien i combinar-les podria assignar a un jugador estadístiques d'un equip en el qual ja no juga— limita la mostra disponible a com a màxim dos partits per enfrontament en fase regular, reduint la fiabilitat del coeficient α per a jugadors amb poca experiència contra un rival concret.

9 Treball futur

Com a extensions naturals d'aquest treball, es proposa incorporar variables contextuals a la funció de fitness —com el rol posicional o les sinergies observades entre parelles de jugadors—, ampliar l'anàlisi a més temporades i competicions europees, i explorar models de Machine Learning supervisats que permetin predir el rendiment *clutch* a partir de les característiques individuals dels jugadors.

Ús d'eines d'intel·ligència artificial

Durant el desenvolupament d'aquest treball s'han utilitzat eines d'intel·ligència artificial generativa, concretament Claude (Anthropic), com a suport en la implementació del codi. L'ús s'ha limitat a la generació i depuració de fragments de codi Python i Jupyter —incloent el pipeline de dades i l'algorisme genètic— a la consulta sobre l'ús de biblioteques específiques, i a la revisió i millora de la redacció de determinats apartats del document.

Agraïments

A l'Ana, per fer-me veure que unir la meva professió i la meva passió és possible, i per confiar en mi en tot moment. A la meva família, pel suport i la comprensió incondicionals durant aquests quatre anys. Al Guillem, per creure en mi quan més ho necessitava i per fer-me estimar el bàsquet dia a dia. I als meus companys i amics, l'Albert, l'Adrià i el David, per fer-me gaudir d'aquesta etapa i per ser el millor equip que podria haver tingut.

Bibliografia

[1] Dell'Isola, J. (2021). *A Quantitative Approach to Measuring Clutch Performance in the NBA*. Bryant University. [Online]. Available: https://digitalcommons.bryant.edu/honors_cis/9.

[2] G. Giasemidis, "Euroleague API," GitHub, 26 02 2026. [Online]. Available: https://github.com/giasemidis/euroleague_api/tree/main/docs/euroleague_api.

[3] M. Lewis, *Moneyball: The Art of Winning an Unfair Game*,

New York: W. W. Norton & Company, 2004.

[4] X. Xiao et al., "Sports Digitalization: A Review and A Research Agenda," in *ICIS 2017 Proceedings*, Seoul, 2017.

[5] V. Sarlis and C. Tjortjis, "Evaluating the effectiveness of machine learning models for performance forecasting in basketball: a comparative study," *Knowledge and Information Systems*, Springer, Mar. 2024. [Online]. Available: https://link.springer.com/article/10.1007/s10115-024-02092-9

[6] Y. Lim, J. Park, and S. Kim, "Enhancing Basketball Game Outcome Prediction through Fused Graph Convolutional Networks and Random Forest Algorithm," *PMC*, 2023. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC10217531/

[7] F. Thabtah, L. Zhang, and N. Abdelhamid, "Predictive Analysis of Successful Basketball Shots: The Euroleague Case," ResearchGate, 2022. [Online]. Available: https://www.researchgate.net/publication/327703497

[8] E. Štrumbelj et al., "Trends in NBA and Euroleague basketball: Analysis and comparison of statistical data from 2000 to 2017," *PLOS One*, 2019. [Online]. Available: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0223524

[9] S. Plakias, C. Kokkotis, D. Pantazis, and T. Tsatalas, "Comparative analysis of key performance indicators in Euroleague and national basketball leagues," ResearchGate, 2024. [Online]. Available: https://www.researchgate.net/publication/381854156**.**

[10] G. Vinué, "A Basketball Big Data Platform for Box Score and Play-by-Play Data," *Big Data*, 2024. [Online]. Available: https://journals.sagepub.com/doi/abs/10.1089/big.2023.0177

[11] Y. Solomonov, I. Avugos, and M. Bar-Eli, "Do clutch players win the game? Testing the validity of the clutch player's reputation in basketball," *Psychology of Sport and Exercise*, vol. 16, pp. 130–138, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S1469029214001459.

[12] C. Cao, J. Price, and D. Wolfers, "Performance Under Pressure in the NBA," *Journal of Sports Economics*, vol. 12, no. 3, pp. 231–252, 2011. [Online]. Available: http://media.clemson.edu/economics/gated/data/golf/Performance%20Under%20Pressure%20NBA%20Journal%20of%20Sports%20Economics-2011-Zheng%20Cao-231-52.pdf.

[13] V. Sarlis, D. Chatziilias, C. Tjortjis, and D. Mandalidis, "A Data Science and Sports Analytics Approach to Decode Clutch Dynamics in the Last Minutes of NBA Games," *Machine Learning and Knowledge Extraction*, vol. 6, no. 3, pp. 2062–2083, 2024. [Online]. Available: https://www.mdpi.com/2504-4990/6/3/102.

[14] "NBA Plus-Minus & Impact Metrics in Basketball Explained," NBASstuffer, 2017. [Online]. Available: https://www.nbastuffer.com/analytics101/nba-plus-minus/.

[15] L. Grassetti, R. Bellio, L. Di Gaspero, et al., "An extended regularized adjusted plus-minus analysis for lineup management in basketball using play-by-play data," *IMA Journal of Management Mathematics*, vol. 32, no. 4, pp. 385–409, 2021. [Online]. Available: https://www.researchgate.net/publication/345490672.

[16] A. Luque, J. M. Palomares, A. Giráldez, and J. C. Riquelme, "Players' selection for basketball teams, through Performance Index Rating, using multiobjective evolutionary algorithms," *PLOS One*, vol. 14, no. 9, 2019. [Online]. Available: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0221258

[17] J. H. Holland, *Adaptation in Natural and Artificial Systems*, Ann Arbor: University of Michigan Press, 1975.

[18] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Reading, MA: Addison-Wesley, 1989.

ANNEX

A1. DIAGRAMA DE GANTT

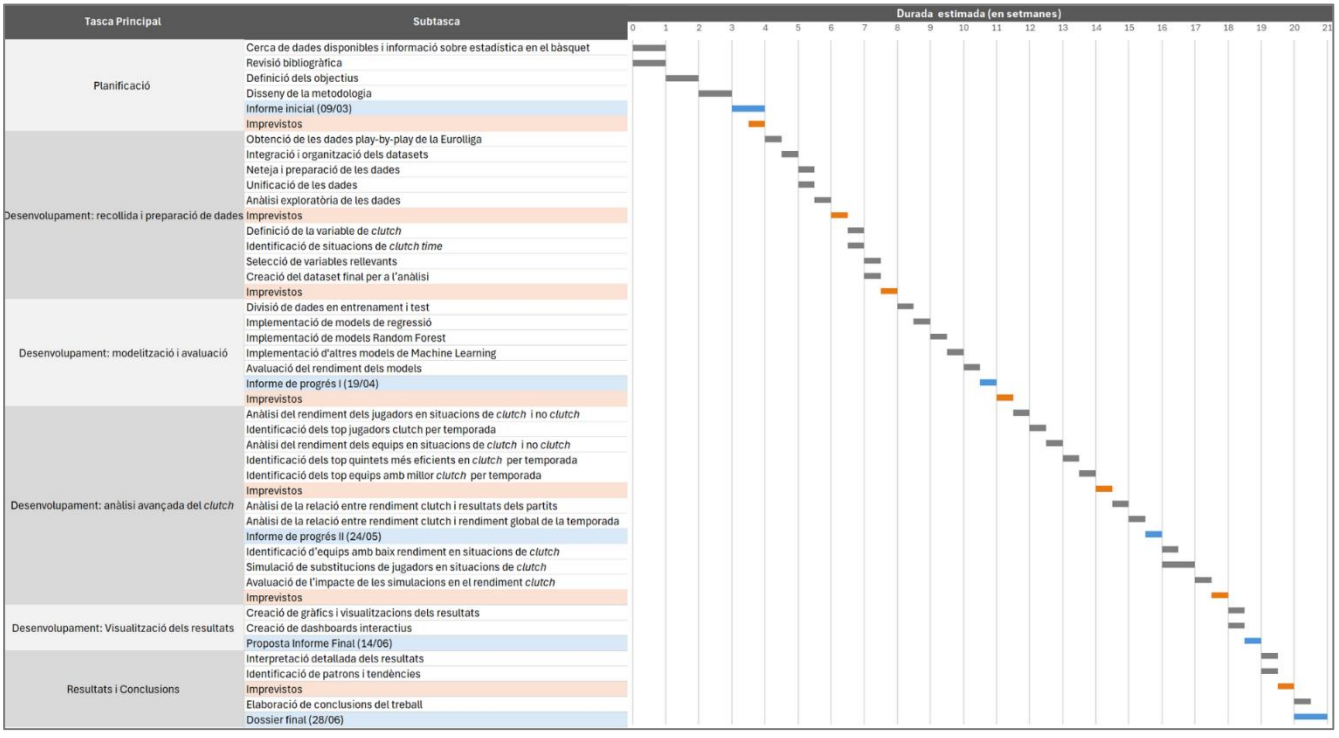

Figura A1.1. Cronograma del projecte

A2. PIPELINE DE LES DADES

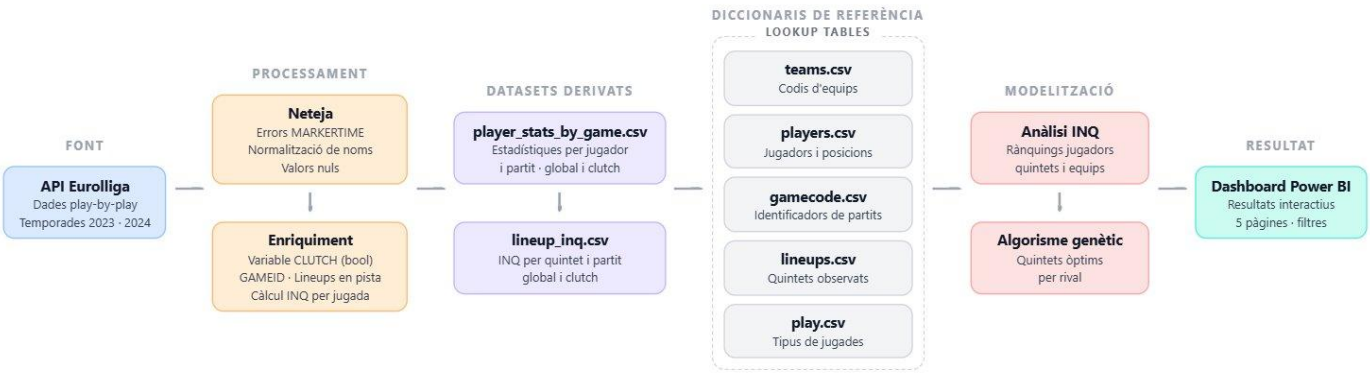

Figura A2.1. Pipeline de les dades

A3. MODEL RELACIONAL DE DADES

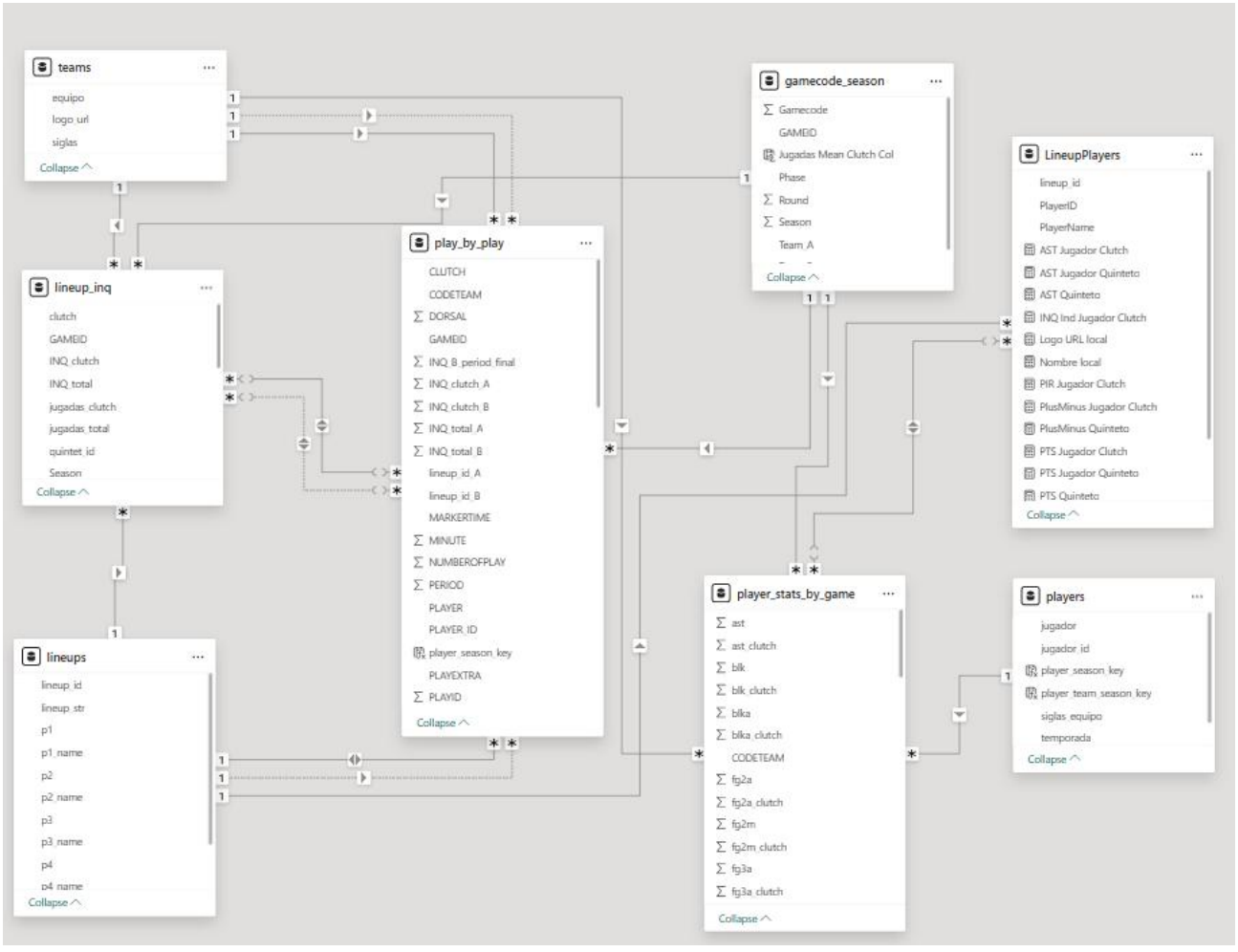

Figura A3.1. Model relacional de les dades

A4. PIPELINE DE L'ALGORISME GENÈTIC

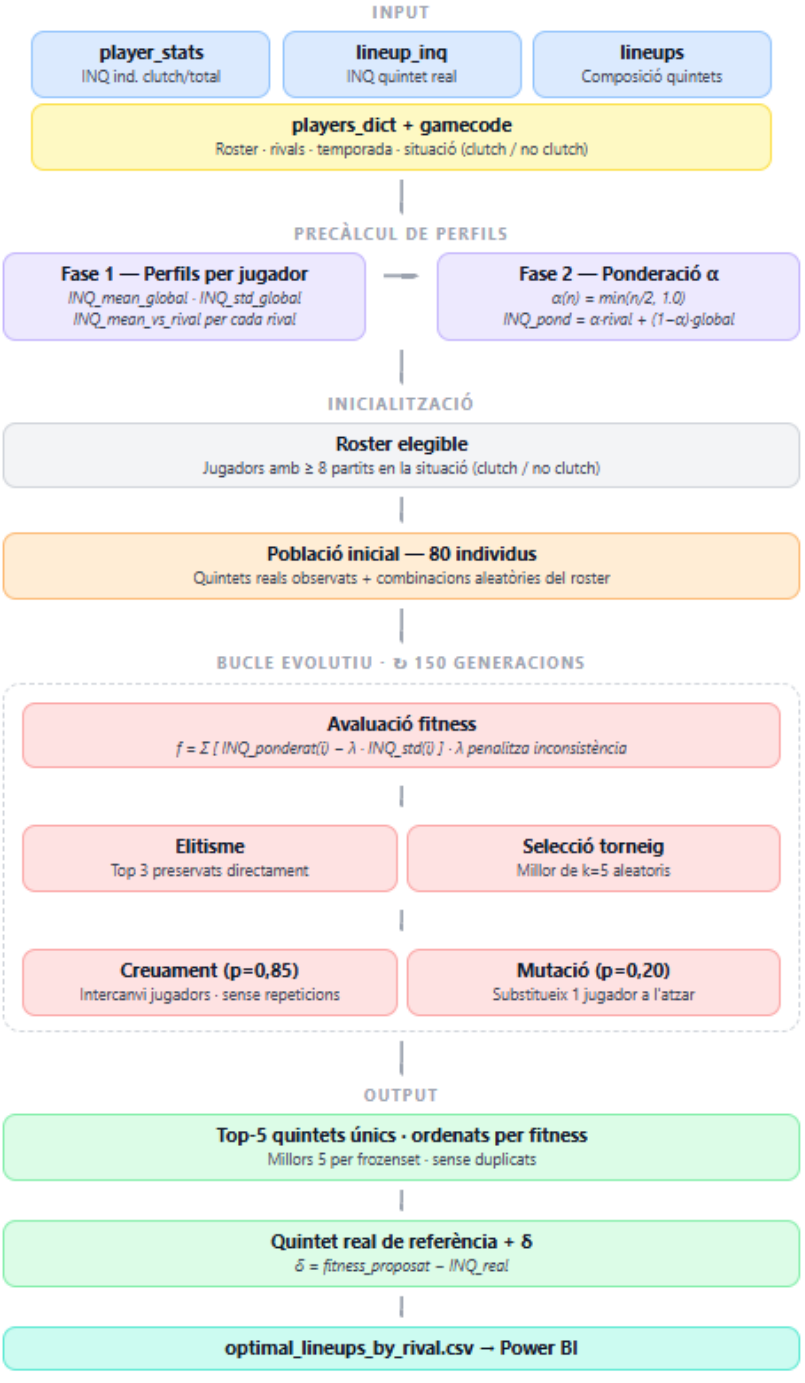

Figura A4.1. Pipeline de l'algorisme genètic

A5. ANÀLISI DE REPRESENTATIVITAT ESTADÍSTICA DELS FILTRES DE QUALITAT

	Context	Filtre	N retinguts	% retinguts	Justificació
0	Quintets global	≥2 partits (~P70)	5593	46.6%	La distribució de partits és molt concentrada en valors baixos. S'aplica un mínim metodològic de 2 partits (~P70) per garantir que el rendiment no correspon a un únic encontre puntual.
1	Quintets clutch	≥2 partits clutch (mínim metodològic)	514	22.2%	La distribució clutch és molt concentrada en 1 partit. S'aplica mínim de 2 partits com a criteri metodològic.
2	Jugadors global	≥24 partits (P40)	361	61.9%	Percentil 40 de la distribució de partits sobre tots els jugadors.
3	Jugadors clutch	≥8 partits clutch (P40)	303	60.2%	Percentil 40 de la distribució de partits clutch sobre jugadors amb almenys 1 partit clutch.

```
NOTA: No s'aplica filtre de jugades/partit als quintets perquè la ponderació
per jugades incorporada al càlcul de l'INQ ja garanteix que els quintets amb
poca activitat tinguin un pes proporcional al seu volum real d'activitat.
```

Figura A5.1.Resum dels filtres de representatitvat estadística aplicats amb justificació

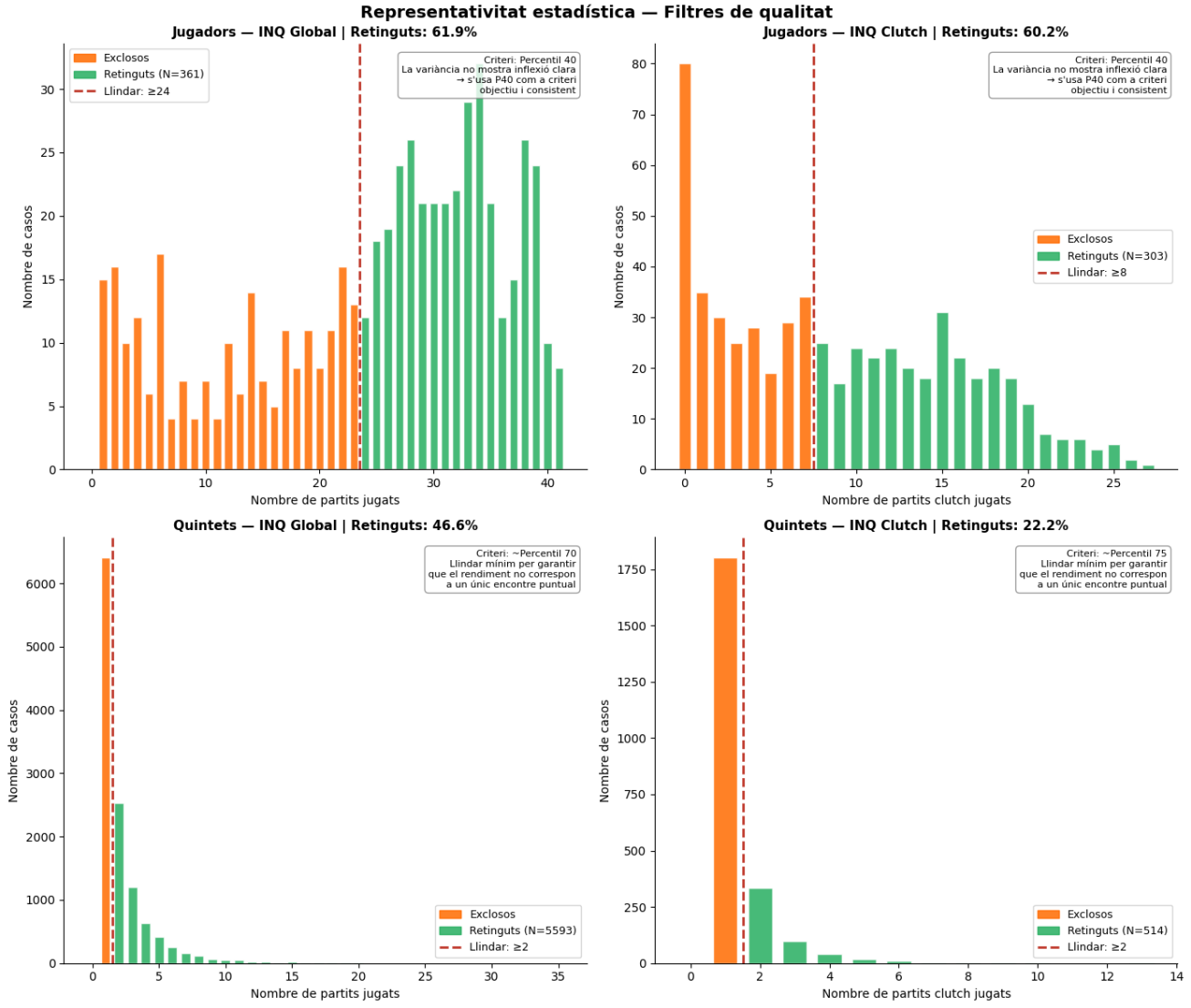

Figura A5.2.Representativitat estadística dels jugadors I quintets segons la situació de joc

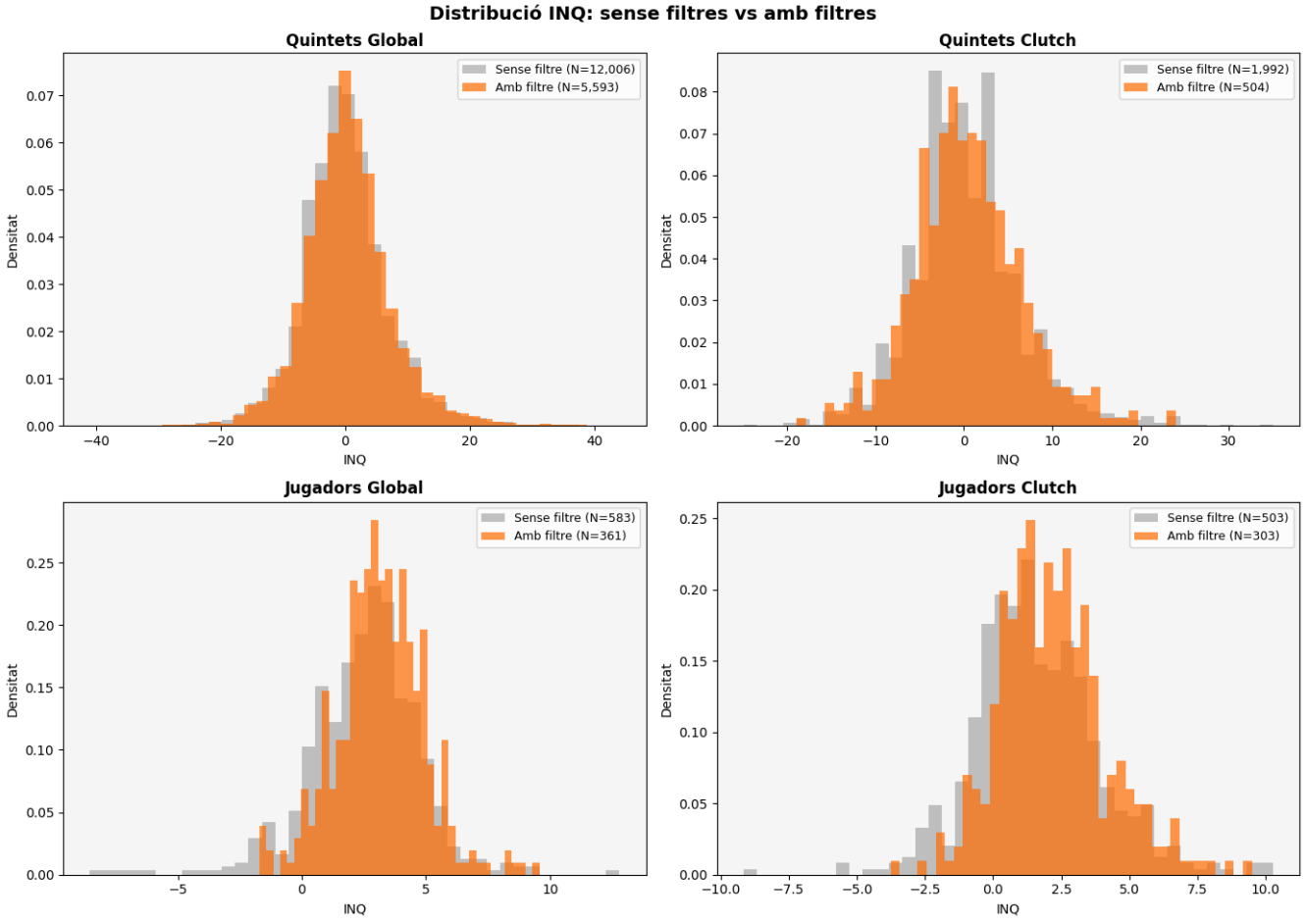

Figura A5.3.Distribució de l'INQ amb i sense filtres aplicats

A6. DASHBOARD DE POWERBI

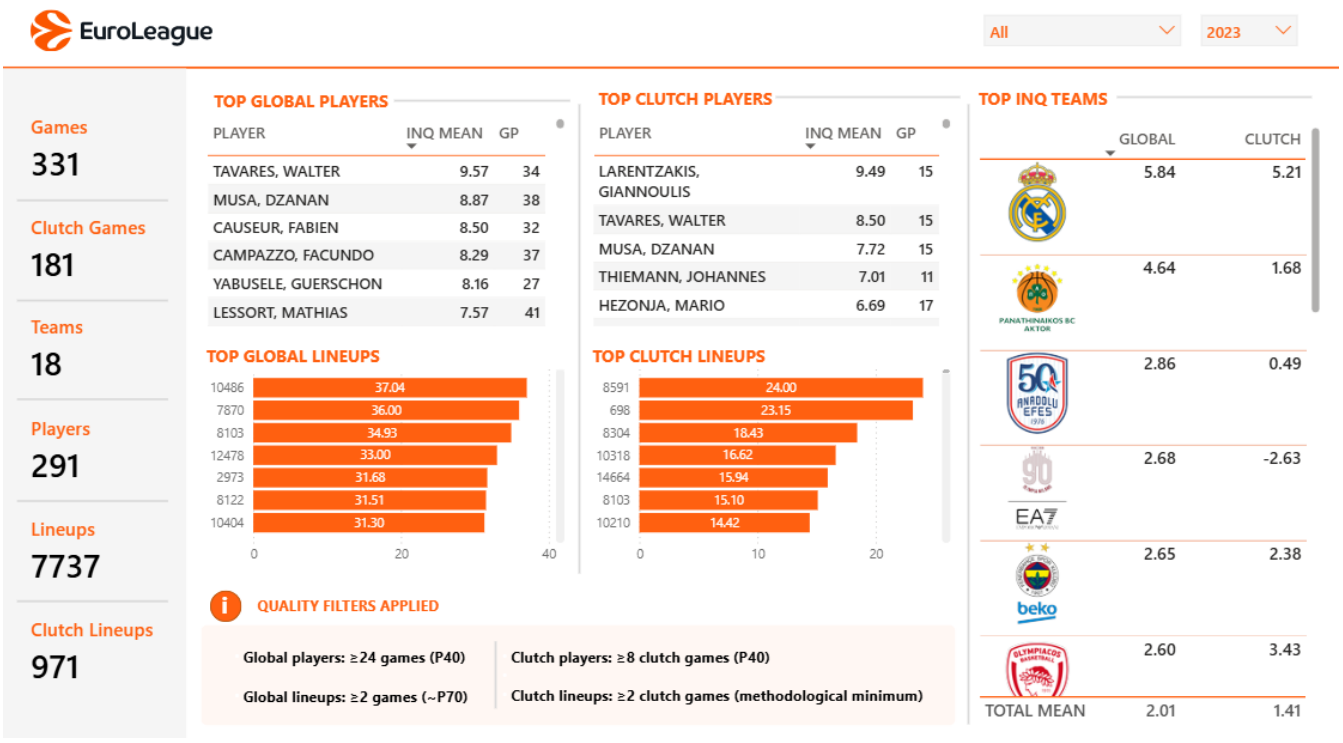

Figura A6.1. Pàgina de rànquings de rendiment (jugadors, quintets i equips) de la temporada 2023



Figura A6.2. Pàgina de rànquings de rendiment (jugadors, quintets i equips) de la temporada 2024

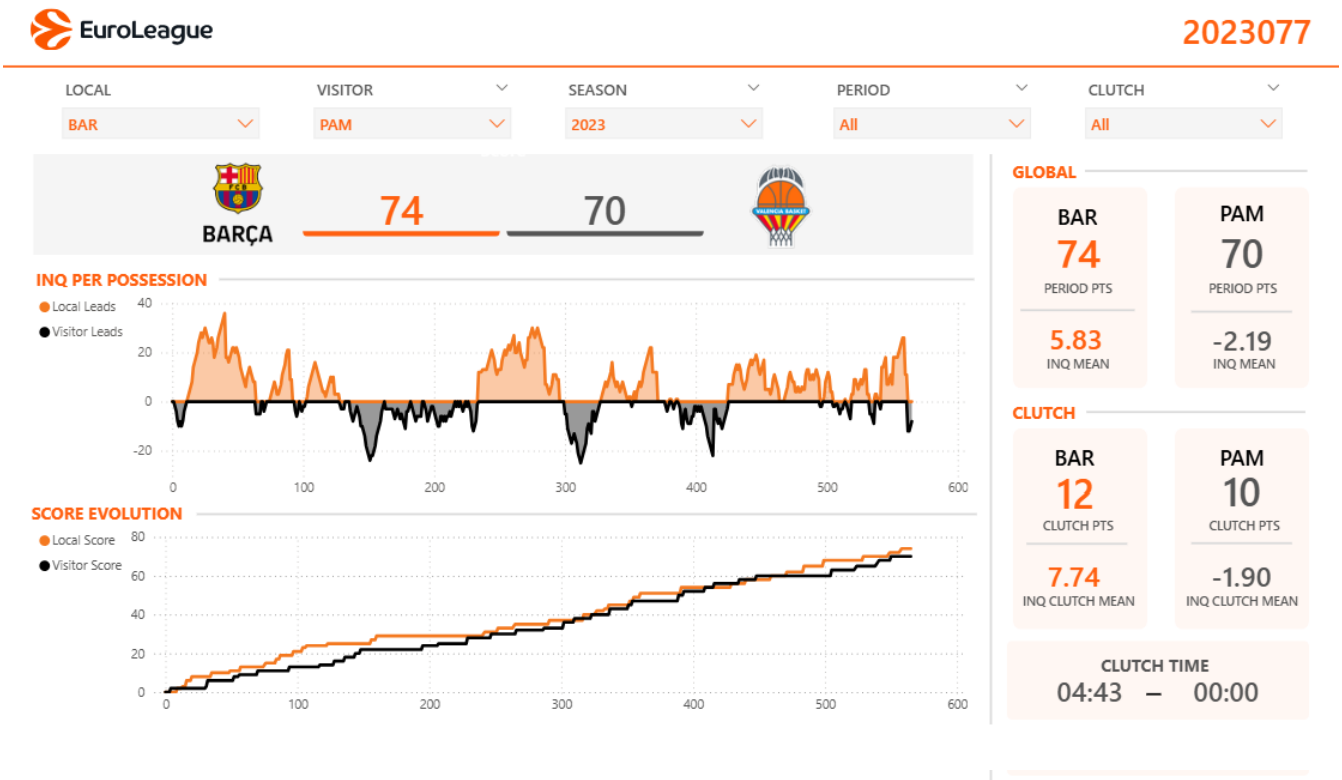

Figura A6.3. Pàgina d'evolució del partit amb visualització del flux de l'INQ al llarg del temps.

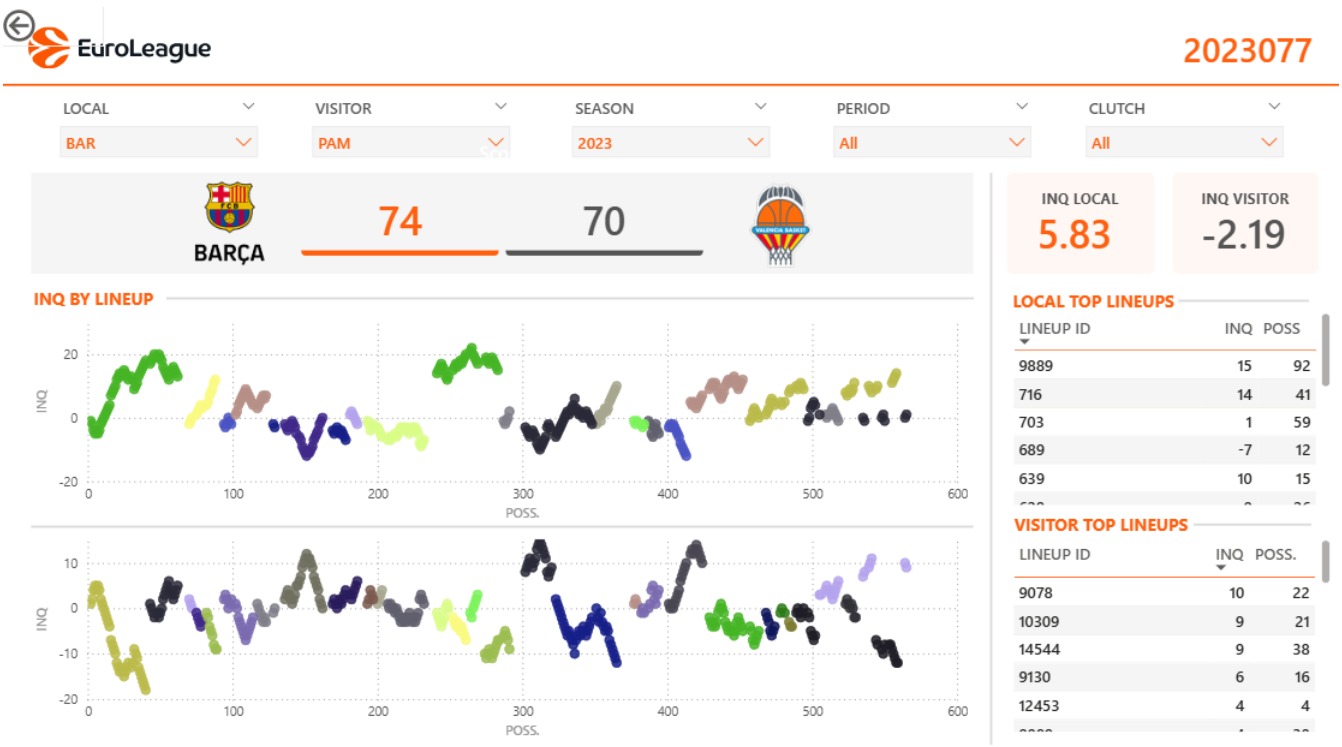

Figura A6.4. Pàgina d'anàlisi de partit, amb l'INQ per quintet al llarg de les possessions per als equips local i visitant

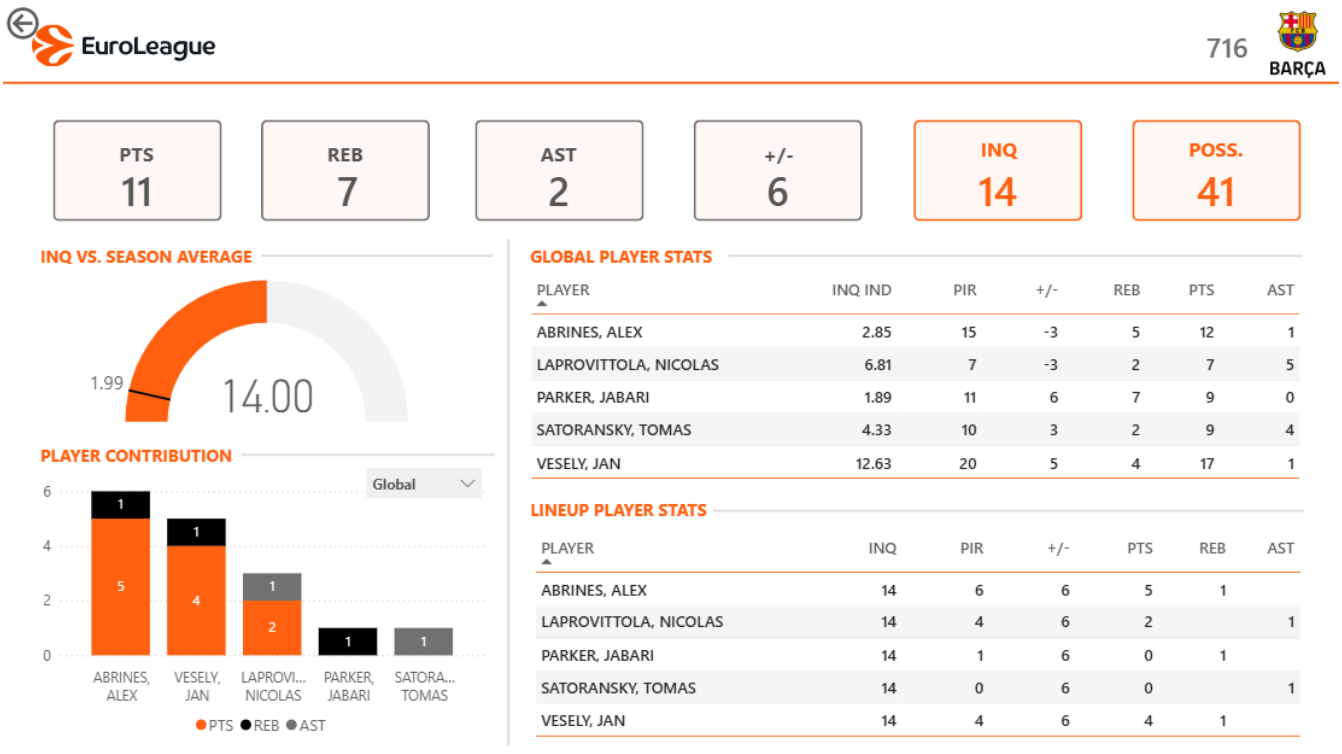

Figura A6.5. Pàgina de detall de quintet, amb les estadístiques individuals dels jugadors i la seva contribució al rendiment col·lectiu del quintet en el partit seleccionat.



Figura A6.6. Pàgina de recomanació de quintets, amb el top-5 proposat per l'algorisme genètic i la comparació amb el quintet real de referència

A7. ANÀLISI COMPARATIVA INQ VS PIR

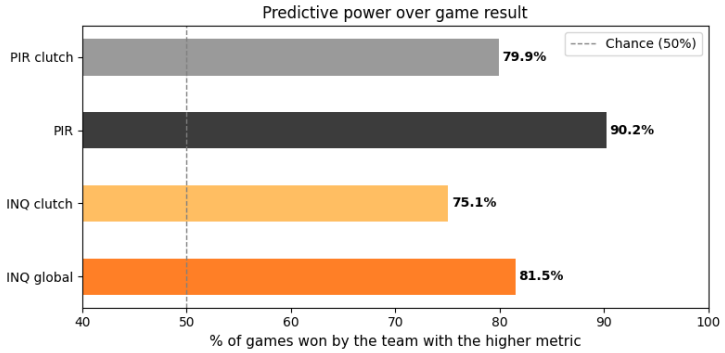

Figura A7.1. Gràfic de barres horitzontals amb el % de partits guanyats per l'equip amb la mètrica superior per a INQ-global, INQ-clutch, PIR i PIR-clutch.

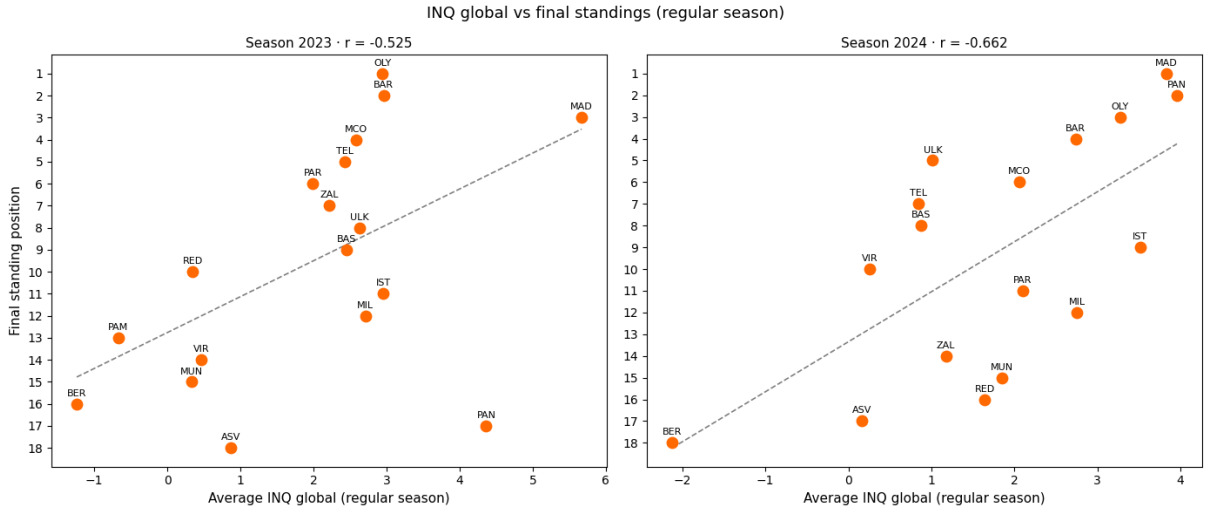

Figura A7.2. Scatter plot amb un punt per equip, mostrant l'INQ-global mitjà a l'eix X i la posició a la classificació a l'eix Y, amb la línia de regressió. Un gràfic per a cada temporada (2023 i 2024)

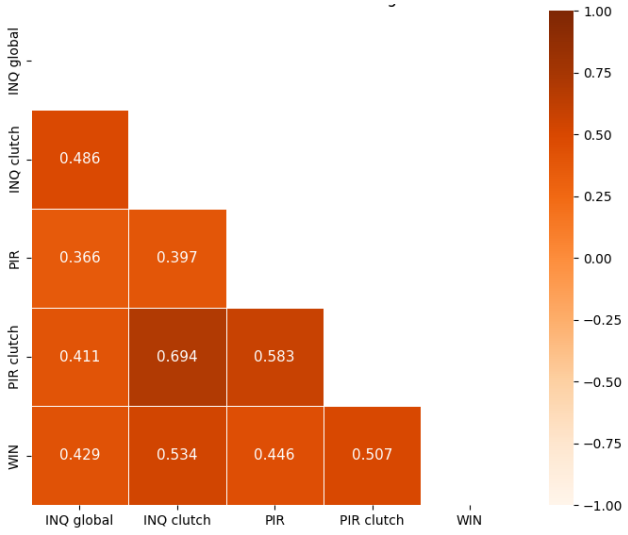

Figura A7.3. Heatmap de la matriu de correlació entre INQ-global, INQ-clutch, PIR, PIR-clutch i WIN.