

# Grau en Estadística

---

**Títol:** L'ús i l'anàlisi del Big Data ha contribuït en la modificació del llançament a cistella en el bàsquet NBA?

*Anàlisi estadística de l'evolució del llançament a cistella en la NBA del Segle XXI.*

**Autor:** Joan Anglès Escardívol

**Director:** Ernest Pons Fanals

**Departament:** Departament d'Econometria, Estadística i Economia Aplicada



*“He fallado más de 9000 tiros en mi carrera. He perdido casi 300 partidos. 26 veces han confiado en mí para tomar el tiro que ganaba el partido y lo he fallado. He fracasado una y otra vez en mi vida y es por lo que tengo éxito.” (Michael Jordan, 2010)*

## Agraïments.

M'agradaria adreçar els meus agraïments a tothom del meu voltant que m'ha acompanyat i fet costat durant la realització d'aquest TFG com a etapa de tancament.

Sobretot, donar les gràcies més especials a la meva dona que m'ha ajudat a tirar endavant en la decisió de finalitzar aquest treball i que ha dedicat gran part del seu temps a estar amb el nostre fill perquè jo pogués dedicar-me a elaborar aquest TFG.

Agrair al meu tutor, Ernest Pons i Fanals, haver volgut agafar-me com alumne sabent la situació especial que tenia i la seva dedicació, per treure el temps d'on ha pogut per tal d'atendre en el meu treball i les seves acurades revisions i recomanacions.

## Resum.

En el món de l'esport cada cop s'utilitzen més les eines del Big Data per millorar el rendiment dels equips, aquest fet ha comportat que en alguns aspectes del joc s'hagi vist traduït en una evolució i canvi d'estratègies amb el pas dels anys.

Aquest Treball de Final de Grau, es focalitza en el bàsquet, concretament en la NBA, en l'evolució que ha tingut el llançament a cistella al llarg del segle XXI.

De manera introductòria es fa una recopilació de l'Estat de l'Art pel que fa al format i estructura de la NBA i les aplicacions del Big Data en el bàsquet NBA.

L'objectiu d'aquest estudi és comprovar si el canvi de tendència en el llançament a cistella és fruit de l'anàlisi de les dades sobre el joc que els equips han anat recopilant de forma massiva.

Es presenta com s'han extret i treballat les dades pels diferents anàlisis descriptius i estadístics que serviran per donar significativitat als canvis de tendència que s'observen en el llançament a cistella.

Finalment, es presentaran les conclusions a les que s'han arribat i possibles aplicacions dels anàlisis fets per a altres aspectes del bàsquet o per a altres esports.

## Paraules clau.

BÀSQUET, ESTADÍSTICA, ANÀLISIS, BIG DATA, ZONA DE LLANÇAMENT A CISTELLA, GRÀFICS, MODELITZACIÓ LINEAL.

## Abstract.

In the world of sport, Big Data tools are increasingly being used to improve the performance of teams, this fact has meant that some aspects of the game have been modified over the years.

In this Final Degree Project, we will focus on basketball, specifically the NBA, on the evolution of the Basket shot in the 21st century.

As introduction, a compilation of the State of the Art is made regarding the format and structure of the NBA and the applications of Big Data in NBA basketball.

The objective of this study is to verify if the shooting trend change is consequence of the game's Big Data analysis done by the teams.

Subsequently, it is presented how the data has been extracted and worked through the different descriptive and statistical analyzes that will be used to give significance to the changes in trend observed in the Basket shot.

Finally, conclusions will be defined and completed with some other applications of the analyzes made over other Basketball issues or other sports.

## Key words.

BASKETBALL, STATISTICS, ANALYSYS, BIG DATA, BASKETBALL SHOT AREA, GRAPHS, LINEAR MODELING.

## Classificació AMS.

- **62J05** Linear regression.
- **62H20** Measures of association (correlation, canonical correlation, etc.).
- **62H30** Classification and discrimination: clúster analysis.
- **97K40** Descriptive statistics.
- **97N40** Numerical analysis.

# Índex de figures.

## Figures.

Figura 1. Esquema comparatiu de les posicions de llançament a cistella dels equips de la NBA de la temporada 2001-02 i la temporada 2021-22.....	12
Figura 2. Mapa amb els diferents equips NBA amb les seves respectives zones geogràfiques.....	14
Figura 3. Quadre de Playoffs NBA. ....	15
Figura 4. Zones de la pista de bàsquet.....	16
Figura 5. Posicions del jugador de bàsquet a pista. ....	16
Figura 6. Esquema exemple de predicció de resposta dels rivals. ....	18
Figura 7. Esquema de les zones de llançament a cistella a la pista de bàsquet. ....	20
Figura 8. Capçalera del format de presentació de les variables en SAP. ....	22
Figura 9. Exemple de les dades extretes en les 5 primeres variables creades.. ....	22
Figura 10. Canvi de tendència en el llançament de 3 PTS i % de diferència entre temporades .....	25
Figura 11. Percentatge llançaments anotats de 3 punts.....	27
Figura 12. Tirs intentats per zona de llançament. ....	28
Figura 13. Zones de llançaments amb canvi de tendència significativa.....	28
Figura 14. Tirs intentats per posició del jugador a pista.....	29
Figura 15. Zona de llançament a cistella amb canvi de tendència.....	30
Figura 16. Llançaments a cistella de 3 punts de mitjana per les temporades del estudi per equips.....	31
Figura 17. Taula d'error segons el nombre de clústers per determinar el nombre de grups k per clusteritzar.....	32
Figura 18. Mitjana dels valors de les variables utilitzades per a cada clúster obtingut.....	33
Figura 19. Llançaments intentats de 3 punts per cada clúster per temporada.....	33
Figura 20. Llançaments intentats de mitja distància per cada clúster per temporada. ....	34
Figura 21. Llançament de dintre de l'ampolla per a cada clúster per temporada. ....	34
Figura 22. Llançaments totals intentats de mitjana per partit per temporada. ....	35
Figura 23. Llançaments de tirs lliures intentats de Mitjana. ....	36
Figura 24. Percentatge de victòria per cluster. ....	36
Figura 25. Histograma de normalitat dels llançaments de 3 punts intentats. ....	37
Figura 26. Resultats test de Normalitat en R. ....	38
Figura 27. Resultats test de Variàncies entre dues variables. ....	39
Figura 28. Resultats test de mitjanes entre T3_FGA vs Temporada. ....	40
Figura 29. Resultat del test de correlació entre T3_FGA i MR_FGA. ....	41
Figura 30. Mapa de correlacions entre la variable T3_FGA amb la resta de significatives....	42
Figura 31. Resultat de la modelització lineal amb variable resposta llançament de 3 punts intentats.....	42
Figura 32. Mapa de correlacions entre el percentatge de victòria i la resta de variables. ....	43
Figura 33. Resultat de la modelització lineal amb variable resposta percentatge de victòria.....	44

## Taules.

Taula 1. Llistat de variables seleccionades per a l'estudi de les zones de llançament a cistella.....	21
Taula 2. Llistat de variables seleccionades per a l'estudi d'estadístiques del bàsquet.....	22
Taula 3. Variables creades a partir de les de l'extracció.....	23
Taula 4. Tècniques estadístiques utilitzades en l'estudi.....	24
Taula 5. Diferències percentuals entre temporades.....	26
Taula 6. Diferències percentuals zones de llançament mitja distancia vs 3 punts.....	30
Taula 7. Resultats test de Normalitat llançament de 3 punts.....	38
Taula 8. Resultats test de Normalitat de les diferents variables seleccionades.....	38
Taula 9. Test de Mitjanes de les variables rellevants de l'estudi.....	40
Taula 10. Resultats correlacions respecte el llançament de 3 punts.....	41



# Índex.

<b>1. INTRODUCCIÓ</b> .....	<b>10</b>
<b>2. CONTEXT I ESTAT DE L'ART</b> .....	<b>14</b>
2.1. Estructura i organització de la NBA.....	14
2.2. Situació tècnico-tàctica del bàsquet en la NBA.....	15
2.3. Història del llançament de 3 punts .....	17
2.4. Antecedents de l'aplicació de noves eines d'anàlisi de les dades en el bàsquet .....	17
<b>3. METODOLOGIA</b> .....	<b>20</b>
3.1. Extracció de les dades .....	20
3.2. Depuració de les dades .....	23
3.3. Tècniques estadístiques .....	24
<b>4. ANÀLISI DESCRIPTIU DE L'EVOLUCIÓ DE LES ZONES DE LLANÇAMENT EXTERIOR</b> .....	<b>25</b>
4.1. Evolució del llançament de 3 punts per Temporada de mitjana de tots els equips NBA.....	25
4.2. Evolució del percentatge d'encert de 3 punts per Temporada de mitjana de tots els equips NBA.....	26
4.3. Evolució de tir intentat per les diferents zones de llançament per Temporada de mitjana de tots els equips NBA.....	27
4.4. Posició del jugador en l'evolució de llançament de 3 punts per Temporada .....	29
4.5. Estudi de cas: Houston Rockets com a precursor del canvi.....	29
4.6. Possibles conseqüències en el joc degut al canvi de tendència en el tir exterior.....	35
<b>5. ANÀLISI ESTADÍSTIC DE L'EVOLUCIÓ DE LES ZONES DE LLANÇAMENT EXTERIOR</b> .....	<b>37</b>
5.1. Estudi de normalitat de la variable resposta (T3_FGA).....	37
5.2. Test de mitjanes de la variable "Temporada" per comprovar la seva significativitat .....	39
5.3. Correlació entre el llançament de 3 punts vs mitja distancia.....	40
5.4. Model lineal amb variable resposta T3_FGA .....	41
5.5. Model lineal amb variable resposta WIN_ .....	43
<b>6. DISCUSSIÓ DE RESULTATS I CONCLUSIONS</b> .....	<b>45</b>
<b>7. BIBLIOGRAFIA</b> .....	<b>48</b>
<b>8. ANNEX</b> .....	<b>50</b>

## 1. INTRODUCCIÓ

Des de l'entrada al nou mil·lenni, la tecnologia ha anat evolucionant exponencialment i ha fet que en 20 anys les nostres vides hagin canviat substancialment. L'arribada d'internet va fer avançar la globalització, ja evident en els anys 90, i tenir fàcil accés a poder compartir informació entre diferents àmbits. A mesura que la societat avançava tecnològicament, es podia emmagatzemar cada cop més informació a les nostres computadores però sense saber que fer-ne d'ella. Aquest augment d'informació va fer que empreses tecnològiques comencessin a preguntar-se com es podien aprofitar aquella recopilació de dades i quina capacitat haurien de tenir les computadores per assolir l'emmagatzematge de cada cop més informació.

A partir d'aquest paradigma, sorgeix el concepte de "Big Data" que, segons la RAE, es defineix com a un conjunt d'estratègies que ens possibiliten recopilar i analitzar una gran quantitat de dades, en les quals es detectin patrons ocults que facin visible la informació rellevant.

A nivell d'evolució històrica del Big Data, es poden assenyalar un seguit d'impactes que van fer sorgir i evolucionar aquest terme. L'any 2005, sorgeix el concepte Web 2.0 entre els usuaris a internet, una nova generació de llocs web que permetien a les persones col·laborar i compartir informació en línia d'una forma que fins aquell moment no havia estat possible. La capacitat de la Web 2.0 de generar enormes volums de dades va fer que aparegués la necessitat de processar-los d'una forma àgil, ja que de poc servia disposar d'una gran quantitat d'informació si quan s'aconseguien extreure les dades ja no estaven actualitzades. És el creador de la Web 2.0, Roger Mougala, d'O'Reilly Media, qui esmenta per primera vegada el terme "Big Data".

Cinc anys més tard de l'aparició de concepte Big Data, les dades que es generaven en dos dies eren equivalents a la quantitat de dades generades des de l'inici de la civilització fins al 2003, (E.Schmidt - director executiu de Google, 2010).

Al 2014, els mòbils superen als ordinadors en accessos a Internet. La connexió quasi continua contribueix a generar moltes més dades i millorar la connectivitat amb altres dispositius. L'accés i l'intercanvi de dades arriben a la població de forma massiva al voltant de l'any 2017. La població controla els seus patrons de descans amb polseres, pot saber en que es gasta els diners en aplicacions mòbils i s'informa sobre la possessió de pilota del seu equip de futbol. Les dades estan per totes parts i la població ja està predisposada a utilitzar-la. (Grupo IGN, 2017)

L'impacte de l'ús del Big Data s'ha vist reflectit en diversos àmbits del nostre entorn. L'estudi de grans quantitats de dades ha servit per intentar optimitzar i millorar processos i sistemes en sectors com l'econòmic, el sanitari o el judicial, entre d'altres. El present projecte, planteja el seu anàlisi en un cas específic del sector esportiu.

### L'esport evoluciona a partir del estudi analític

El primer esport i gerent que va decidir apostar per aplicar models matemàtics va ser Billy Beane amb els seus Oakland Athletics, equip de beisbol de la MLB dels EUA. Amb el penúltim pressupost de la lliga va decidir començar a aplicar models matemàtics per tal de captar jugadors i poder compensar la plantilla. Els bons resultats adquirits a partir d'aquesta pràctica, han fet que el beisbol sigui un dels esports on, actualment, més s'estan aplicant les tècniques d'anàlisi.

Després d'aquest fet i conjuntament amb l'evolució ja esmentada del Big Data, la incidència en el món esportiu es va accelerar i a dia d'avui, un 97% d'equips en la Major

League Baseball utilitzen analítics i consultors professionals, a la NFL un 60 % i a la NBA fins a un 80% (Deloitte, 2022).

A continuació s'esmenten els principals objectius que té l'aplicació del Big Data en l'àmbit esportiu (Sports Data Campus,2022):

- **Conèixer-se i conèixer el rival.** Es presenten múltiples opcions relacionades essencialment amb l'anàlisi del propi equip i del rival, entorn competitiu. En aquest sentit, s'utilitzen múltiples mètriques, unes més personalitzades i d'altres menys, que serveixen per descriure de forma objectiva models de joc, sistemes, ocupació d'espais i característiques de plantilles i jugadors.
- **Evitar lesions.** Es recopilen dades esportives, biomètriques, físiques, genètiques i nutricionals. Models d'entrenament amb una gestió de càrrega personalitzada amb l'objectiu de prevenir lesions.
- **Captació de jugadors.** Si s'ha de substituir un jugador per un altre, cercar el jugador amb característiques similars per reduir el marge d'error. Utilitzant les variables adequades trobar el més semblant.

Aquest canvi de paradigma derivat de l'ús del Big Data ha contribuït a que els equips comencin a introduir nous conceptes i tàctiques per millorar el joc. A partir d'aquí es planteja la següent hipòtesi: Aquests canvis han contribuït a modificar el propi joc? Aquests equips que han aplicat els canvis derivats del estudi de dades massives, ha significat un augment en la probabilitat de victòria?

Per tal de contestar aquestes qüestions, el present treball es centra en la recerca i estudi del cas al bàsquet i concretament a l'entorn de la NBA.

## La NBA entra en l'estudi del Big Data

La NBA és la lliga de bàsquet més popular del món tant en aspectes econòmics com esportius.

A partir del 2010 alguns equips van començar a introduir elements d'anàlisi en les dades que havien començat a observar en altres esports. Un exemple d'això és Daryl Morey, el director general dels Houston Rockets, un equip que va canviar completament la seva forma de jugar un cop Morey es va incorporar a la plantilla. Morey és un expert informàtic que va arribar als Houston Rockets l'any 2006. Quan se'l va contractar l'equip buscava algú que fos capaç de veure els jugadors des d'una altra perspectiva. Aquell mateix any, Morey va fundar la MIT Sloan Sports Analytic Conference, un cicle de conferències que tenen com objectiu parlar de la importància del moviment sabermetrics<sup>1</sup> en l'esport professional, estudiant les possibilitats que ofereix i fins on pot arribar. Així Morey va contribuir en què, els següents anys, les diferents franquícies de bàsquet invertissin molt més en tenir professionals en l'anàlisi de dades. Es va introduir: l'anàlisi de les posicions de llançament, relacions entre els aspectes morfològics dels jugadors respecte les seves classificacions ofensives/defensives per determinar el seu potencial, predicció de percentatge de victòria segons les variables estadístiques del joc, predicció de la substitució d'un jugador a partir de trobar aquell més semblant a nivell numèric, entre d'altres.

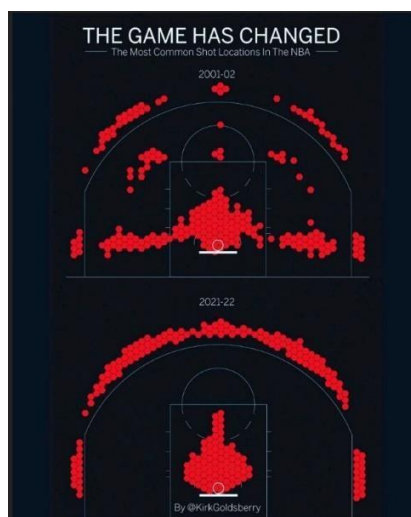
---

<sup>1</sup>Sabermetrics: Recollida i recopilació de dades rellevants que es produeixen en els partits, per tal de respondre preguntes específiques. (James & Wirth, 1986).

L'arribada de Morey va suposar un canvi significatiu pels Houston Rockets i el bàsquet en general. Convençut d'integrar l'anàlisi de dades en l'esport, va ser dels primers en convèncer a la franquícia d'instal·lar un sistema de localització de vídeo, amb el qual podia extreure dades directament del propi terreny de joc. Gràcies a això, l'equip va arribar a la conclusió que el llançament de tres tenia més importància que el de dos a l'hora d'augmentar la probabilitat de victòria. Això va comportar a un canvi en el sistema de joc de l'equip amb l'increment dels llançaments de tres punts. En els següents anys els Houston Rockets van obtenir bons resultats i molts equips van començar a invertir també en analitzar i modificar el seu joc.

Tant és així, que el següent esquema de tir de la Figura 1, sembla corroborar que aquesta introducció va modificar el comportament dels equips, i en conclusió els llançaments de tres punts (fora del segon semicercle) van passar a ser prioritaris respecte els de dos punts (espai dintre el semicercle i fora de l'ampolla).

És Daryl Morey i els seus Houston Rockets els que inicien aquesta tendència o hi ha més equips involucrats? En cas afirmatiu, són els que més han invertit en anàlisi de dades en les seves franquícies? O bé, és un o més jugadors els que van contribuir a aquest canvi?



*Figura 1. Esquema comparatiu de les posicions d'lançament a cistella dels equips de la NBA de la temporada 2001-02 i la temporada 2021-22.*

Una part d'aquest treball es focalitza en l'estudi de l'evolució de les posicions de llançament a cistella més comunes en els equips de la NBA entre les temporades 2001-02 i 2021-22. Amb aquest estudi es vol contrastar si el canvi de tendència en el joc ve donat per les conclusions extretes de la recopilació i anàlisi de dades.

Per donar resposta a aquestes preguntes, s'han plantejat els següents objectius:

O1. Estudiar la història en l'evolució de l'aplicació de l'anàlisi de dades massives en els equips de l'NBA. (Quins han estat pioners, quins han estat els principals factors estudiats, com ha afectat en l'evolució del joc, etc.).

O2. Analitzar descriptivament i analíticament el canvi de tendència del llançament a cistella i possibles conseqüències derivades.

O3. Analitzar la possible correlació entre el canvi de posicions de llançament a cistella i la inversió en anàlisi de dades massives a la NBA.

O4. Investigar els factors decisius en el canvi de la tendència de tir i el seu efecte en l'increment de victòries.

O5. Obtenir una predicció en probabilitat de victòria utilitzant un Model estadístic, aquesta equació resultant que dependria de diferents variables.

## 2. CONTEXT I ESTAT DE L'ART

### 2.1. Estructura i organització de la NBA

La NBA ("National Basketball Association") és la lliga professional de bàsquet més important a nivell econòmic, competitiu i cultural del món. La lliga es juga als Estats Units, encara que hi han partits esporàdics que es juguen a altres parts del món per promocionar la pròpia lliga.

La NBA es fundà a Nova York al 1946, creada pels propietaris dels principals pavellons esportius del nord-est i mig-oest dels Estats Units. La lliga buscava que a cada ciutat important del país hi hagués una franquícia amb un equip que la representi (Pàgina oficial NBA Espanya, 2021). Es va iniciar amb 17 equips en diferents ciutats. A mesura que van passar els anys, el creixement de la lliga s'estancava i el bàsquet seguia sent un esport principalment d'universitats, fins que l'any 1979 amb l'arribada de dos jugadors: Larry Bird i Magic Johnson a dos equips amb una rivalitat històrica com eren els *Boston Celtics* i *Los Angeles Lakers*, van fer que la lliga comencés a créixer de forma exponencial (Pàgina oficial NBA Espanya, 2021). La rivalitat entre aquests dos equips, que es van trobar en diferents finals de la competició, va fer que la lliga anés creixent fins als 30 equips que la formen a dia d'avui.

Com que els Estats Units és un país amb una gran extensió, la lliga va dividir la competició en dos conferències: la *Conferència Est* i la *Conferència Oest* que es subdivideixen en 3 divisions cadascuna:

#### CONFERENCIA EST:

- **Divisió Atlàntica:** Toronto Raptors, Boston Celtics, Philadelphia 76ers, Brooklyn Nets y New York Knicks.
- **Divisió Central:** Milwaukee Bucks, Indiana Pacers, Chicago Bulls, Detroit Pistons y Cleveland Cavaliers.
- **Divisió Sur-Est:** Miami Heat, Orlando Magic, Washington Wizards, Charlotte Hornets y Atlanta Hawks.

#### CONFERENCIA OEST:

- **Divisió Nord-Oest:** Denver Nuggets, Utah Jazz, Oklahoma City Thunder, Portland Trail Blazers y Minnesota Timberwolves.
- **Divisió Pacífica:** Los Angeles Lakers, Los Angeles Clippers, Sacramento Kings, Phoenix Suns y Golden State Warriors.
- **Divisió Sur-Oest:** Houston Rockets, Dallas Mavericks, Memphis Grizzlies, New Orleans Pelicans y San Antonio Spurs.



Figura 2. Mapa amb els diferents equips NBA amb les seves respectives zones geogràfiques. (Basket World, 2022)

Un dels principals elements característics de la NBA és la igualtat entre franquícies, d'aquí la creació d'un *Draft* a l'inici de cada temporada on aquells equips que han quedat pitjor classificats, tenen més possibilitats d'escollir els millors jugadors joves que entren a la lliga per a la següent temporada per tal d'intentar equilibrar la competició.

El format de la competició segueix el següent ordre:

- **Draft NBA:** Les franquícies segons l'ordre que els hagi tocat en funció de la classificació de l'any anterior, seleccionen els jugadors joves que incorporaran a la plantilla.
- **Pretemporada:** els equips juguen una sèrie de partits per a preparar als equips per a l'inici de la competició.
- **Lliga Regular:** els equips juguen un total de 82 partits. Els equips de la mateixa Conferència jugaran 4 partits entre ells i contra els de diferent Conferència 2 partits.
- **Fase final – Playoffs:** Els 8 millors classificats (millor percentatge de victòria) de cada Conferència s'enfronten entre ells en eliminatòries de format directe (el que guanya passa a la següent ronda, el perdedor queda eliminat definitivament de la competició) al millor de 7 partits.
- **Final NBA:** El guanyador dels *Playoffs* de la Conferència Est contra el guanyador dels *Playoffs* de la Conferència Oest.

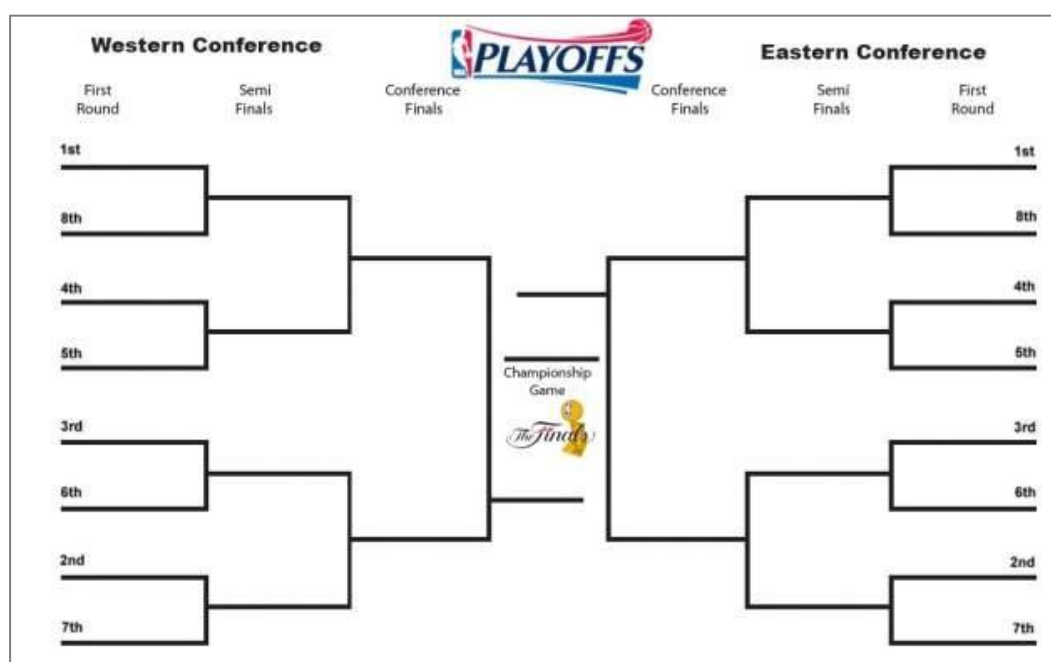


Figura 3. Quadre de Playoffs NBA (Basket World, 2022)

## 2.2. Situació tècnico-tàctica del bàsquet en la NBA

El bàsquet és un esport entre dos equips de cinc jugadors que el seu objectiu és encistellar la pilota a la cistella contrària defensada per l'equip contrari. L'equip que aconsegueix-hi més punts guanya. Per tal de contextualitzar les dades utilitzades en l'anàlisi estadístic d'aquest treball, a continuació es presenten les diferents possibles puntuacions dels tirs anotats a cistella depenent de la zona de llançament:

- **1 punt:** Cada cistella anotada des d'aquesta zona després d'haver rebut una falta, el llançament s'efectua sense cap defensa del contrari.

- **2 punts:** Cada cistella anotada dintre de la zona.
- **3 punts:** Cada cistella anotada fora de la zona.

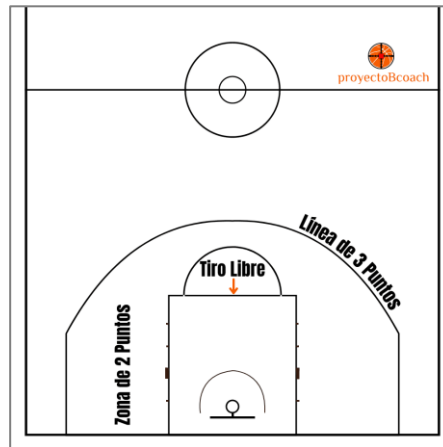


Figura 4. Zones de la pista de bàsquet (Basket World, 2022)

Els 5 jugadors de cada equip han de repartir-se per la zona de joc, d'aquesta forma es defineixen diferents posicions que tindran característiques diferents per a cada jugador. En la NBA es divideixen les posicions de joc i tir en 3:

- **Guard:**
  - Point Guard:** és el jugador que puja la pilota i dirigeix el joc. Es defineix com l'entrenador a pista, solen ser jugadors amb menys alçada en comparació amb les altres posicions, amb molt domini de la pilota, el "pase" i del tir.
  - Shooting Guard:** jugador que la seva funció és centra en la defensa i anotació des de fora de l'ampolla.
- **Forward:**
  - Small-Power Forward:** Jugadors més alts que els *guards* i amb més envergadura, les seves tasques solen ser més físiques ( control de rebot, penetració a cistella, etc.).
- **Center:**
  - Són els jugadors més alts de l'equip, juguen a prop de la cistella. La seva funció es centra en capturar rebots i en anotar després de moviments de recepció d'esquena o bé de continuació de bloqueigs dels *guards* o *forwards*.

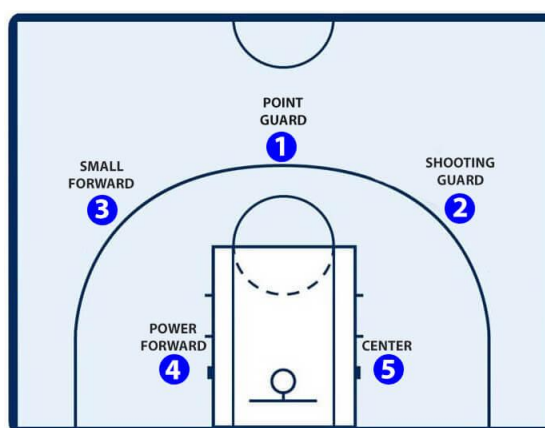


Figura 5. Posicions del jugador de bàsquet a pista (Basket World, 2022)



### 2.3. Història del llançament de 3 punts

A l'inici de la NBA no existia la línia marcada de la zona de 3 punts, o bé s'anotava 1 o 2 punts. Per tal de contrarestar el joc ofensiu dels jugadors més alts, es va decidir provar en competicions universitàries de posar una línia a 6,25m de la cistella per tal de que aquelles cistelles que fossin de més enllà valguessin 3 punts, (Pàgina oficial NBA España, 2021).

L'any 1979 la NBA accepta aquesta prova en les lligues universitàries i l'introdueix a la seva lliga. Aquest fet va modificar el propi joc ja que va fer ampliar els espais de les posicions de joc dels jugadors. Per exemple, en el bàsquet FIBA (Federació Internacional de Bàsquet) no s'introdueix aquesta modificació fins el 1984, (Pàgina oficial NBA España, 2021).

Durant els primers anys, el joc no es va veure molt condicionat, es començava a anotar més per partit però ja venia essent una tendència degut a que els jugadors començaven a estar més preparats gràcies a la professionalització i creixement de la NBA. Els equips encara no tenien especialistes que llancessin des d'aquesta posició, però al llarg dels anys cada cop van anar sorgint més jugadors que es dedicaven exclusivament a anotar des d'aquestes posicions de llançament, (Pàgina oficial NBA España, 2021). Per tant, a mesura que passaven els anys, els equips van començar a fitxar a alguns jugadors especialistes en aquest tipus de tir, però la resta de l'equip seguia prioritant el llançament de 2 punts.

És a partir de la dècada del 2010, que amb l'entrada de les noves tecnologies per estudiar dades massives, algun equip comença a analitzar que si augmenten els llançaments de 3 punts tindran un percentatge d'encert semblant a si ho fessin de 2 punts i aconseguirien un punt més per jugada, (Pàgina oficial NBA España, 2021). D'aquesta forma, no només els especialistes, sinó que la resta de posicions dels jugadors al camp comencen a llançar amb més freqüència de 3 punts, fent que el propi joc a nivell *tècnico-tàctic* comenci a evolucionar cap a una altra direcció.

En el present treball, es vol donar veracitat estadística als factors que han condicionat a aquest augment en el llançament de 3 punts en el bàsquet de la NBA.

### 2.4. Antecedents de l'aplicació de noves eines d'anàlisi de les dades en el bàsquet

Dins de l'era del Big Data, les estadístiques avançades són les més demandades als Estats Units i les franquícies de la NBA estan apostant cada cop més fort per equips de "Data Science" que treballen amb quantitats inimaginables de dades amb la finalitat de millorar el rendiment de l'equip, (Solobasket, revista oficial NBA España, 2018).

El Sports Analytics Conference (SSAC), organitzat pel MIT Sloan, és el congrés per excel·lència que reuneix esport, tecnologia i dades. A continuació, es presenten una selecció d'aplicacions d'eines d'anàlisi de dades que ja s'han introduït en l'entorn de la NBA:

- **El concepte del "GHOSTING"**

Investigació que pretén simular el comportament dels jugadors donat l'inici d'una seqüència utilitzant tècniques d'aprenentatge automàtic (Deep Learning).

Dissenyar una jugada perfecta és una tasca difícil, donat que no depèn només del *scouting*<sup>2</sup> que faci l'equip rival, sinó que també és necessari comptar amb una dosi d'espontaneïtat creativa dels propis jugadors, on l'entrenador s'ha d'anticipar a la defensa que se li plantejarà.

---

<sup>2</sup> Scouting: Exploració i Investigació de como juega el rival.

El “ghosting” el que pretén és ser una eina que, utilitzant una *tablet* a peu de pista, analitzi els possibles escenaris d’una hipotètica jugada, és a dir, que el model d’intel·ligència artificial conegui els patrons defensius específics per a cada equip NBA, i donat aquest esbós temporal d’una jugada, el programa pot adaptar-la al tipus de defensa que rebrà el seu equip.

Els resultats són excel·lents (A. Arbués Sangüesa, Revista Solobasket, 2018), inclús s’arriba a predir el tipus de defensa que farà un equip en molts bloquejos directes. L’estudi inclou una aplicació on l’entrenador pot dibuixar les jugades que vulgui i rebrà les possibles reaccions del equip rival en qüestió.

El “Deep Learning” es basa en bases de dades de milions d’imatges/seqüències, amb aquestes es construeix una xarxa neuronal que funciona com a caixa negra, donat que per si mateixa aprèn quines són les característiques més representatives dels ítems en qüestió.

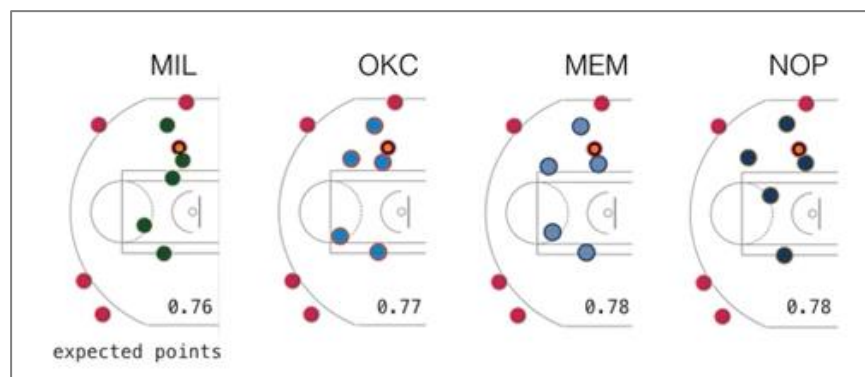


Figura 6. Esquema exemple de predicció de resposta dels rivals

- **Canvi de realitat. “Drafting Errors and Decision Making Bias in the NBA Draft (Daniel Smailofsky)”**

Les tècniques d’intel·ligència artificial també serveixen per replantejar-se la realitat, intentar entendre el perquè de la presa de decisions en el *Draft* o plantejar-se escenaris alternatius com sistemes ofensius de joc.

Escollir un jugador en el *Draft* és una de les tasques més difícils d’un General Manager (GM), ja que si s’encerta la pròpia franquícia podrà créixer amb major facilitat. Els GM segueixen uns patrons a l’hora de prendre una decisió que maximitzi el rendiment de l’equip, però cal assegurar que s’utilitzin els criteris adequats. El sociòleg i entrenador Daniel Smailofsky va investigar els factors estadístics que afectaven a la presa de decisió segons la qual un jugador de la NCAA (Lliga universitària americana de bàsquet) és escollit en el *Draft* i quina correlació amb aquests factors tenien amb el rendiment en la NBA. Els resultats van determinar que la universitat de procedència tenia una alta influència en la posició en la qual quedava seleccionat el jugador, però no tenia correlació amb el rendiment del jugador un cop a la NBA. Per tant, els GM anaven a les millors universitats simplement per la seguretat de que si el jugador després no tenia el rendiment esperat a repercussió mediàtica no seria rellevant, però potser s’estan deixant d’anitzar amb dades empíriques jugadors d’altres zones amb més potencial.

L’estudi determina que les dos estadístiques més influents positivament en el rendiment NBA són el percentatge de rebots i de pèrdues, ambdues estadístiques del joc relacionades amb el control de la pilota (A. Arbués Sangüesa, Revista Solobasket, 2018). També es demostra que els jugadors alts i amb més envergadura surten més afavorits en el *Draft* de la NBA, encara que aquestes característiques no influeixen en el rendiment NBA de forma significativa.

Aquests diferents aspectes que es van demostrant, poden ajudar a les franquícies a millorar en la selecció dels jugadors del *Draft*, i així, a mig termini poden portar a l'èxit.

## El paper de Daryl Morey en la introducció del Big Data a la NBA

El Big Data i l'analítica estan darrere de la forma com es juga a bàsquet d'alt nivell. La intuïció i les sensacions han donat pas al estudi de les dades per guanyar més partits.

Avui dia, els *Center* són capaços d'anotar des de lluny i els *Guard* tenen més envergadura i poden, així, parar a jugadors més alts. Les pròpies posicions de joc tan marcades amb anterioritat s'estan modificant de tal forma que en uns anys tendiran a modificar-se.

El primer en utilitzar la lectura de dades avançades per interpretar-les i modificar el joc del seu equip va ser el GM dels Houston Rockets, Daryl Morey. Aquest enginyer de sistemes i admirador del vicepresident executiu d'operacions de Beisbol Billy Beane i de la seva idea d'incloure l'anàlisi de dades en l'equip de Beisbol Oakland Athletics l'any 2002, va arribar als Houston Rockets al 2007 amb l'idea de que utilitzant les dades podien fer evolucionar el joc i convertir un equip dels últims classificats en un dels equips aspirants al títol.

La seva filosofia en un inici era molt simple: anotar més punts per possessió i aconseguir més possessions extra. Ho remarcava en 4 punts:

1. **Joc ràpid.** Es llança a cistella en els primers segons de possessió.
2. **Més llançaments de triples que de 2 punts.** Un percentatge major al 30% en tirs de 3 punts té molt major benefici que un percentatge del 50 % en tirs de 2 punts.
3. **Més esmaixades<sup>3</sup> i tirs des de l'ampolla.** Els tirs a prop de la cistella tenen un percentatge més elevat.
4. **Tancar el rebot<sup>4</sup>:** aconseguir més rebots ofensius i defensius per tenir noves possessions.

Per aconseguir aquests canvis, es va utilitzar la tecnologia que estava sorgint per mesurar i trackejar<sup>5</sup> (càlcul de coordenades del jugador) els entrenaments i partits, analitzant totes les variables que podien estudiar ( velocitat dels jugadors, salts, nombre de cistelles i encerts de tirs amb o sense defensa, després de bot, etc....).

Les conclusions dels seus estudis va canviar el joc dels Rockets i van acabar donant la raó a Morey. Els Rockets van passar de jugar 93 possessions en atac en la temporada 2007-2008 a 101,5 en la temporada 2017-2018, van passar d'anotar 583 triples (amb un 35,2% d'encert) en una temporada a 1256 triples (amb un 36,2% d'encert). Van passar de 30 victòries a 65 victòries i van acabar sent un clar aspirant a guanyar la NBA (Antoni Daimiel, periodista especialitzat en NBA).

Quan es té una idea clara del tipus de joc que es vol fer, també la franquícia va encaminada a seleccionar o contractar els jugadors amb aquelles característiques que afavoreixin la filosofia marcada per l'equip.

---

<sup>3</sup> Esmaixada: Jugada de bàsquet que consisteix en aconseguir una cistella acompanyant la pilota amb la mà fins a l'anell, introduint-la amb força mitjançant un ràpid moviment de canell.

<sup>4</sup> Rebot: Agafar la pilota després d'un llançament que no ha entrat a cistella.

### 3. METODOLOGIA

#### 3.1. Extracció de les dades

Per resoldre si ha hagut un canvi en la tendència del llançament exterior i, en cas afirmatiu obtenir els factors decisius a aquest canvi, s'ha cercat una font fiable de dades de la *NBA* de la qual s'obtinguin les variables que es consideraran decisives per fer l'estudi.

Entre diferents opcions de bases de dades (*Basketball Real GM, NBN23, STATS*), s'ha decidit treballar amb una font de dades oficial de la *NBA* les dades de la qual estan en format *SAP*. La pàgina web oficial de la *NBA* compta amb una base de dades molt extensa, amb diferents opcions per a treballar-hi i s'ha considerat òptima per al treball donada la gran diversitat de variables quantitatives i qualitatives que ofereix per a poder fer els anàlisis pertinents.

- **Posició de llançament a cistella**

Un dels primers interrogants que es planteja l'estudi és si realment ha hagut un canvi de tendència en el tir exterior, si en l'evolució del període estudiat hi ha un canvi significatiu en el nombre de llançaments de 3 punts. Per tant, s'han extret el nombre de tirs intentats, tirs encertats i percentatge de tir del període estudiat (2001-2022). En la font de dades que es treballa, les variables per obtenir aquests valors de nombre de tirs estan dividits per zones de llançament.

A continuació, es presenta un breu resum d'aquestes zones de llançament i de les variables seleccionades per l'estudi.

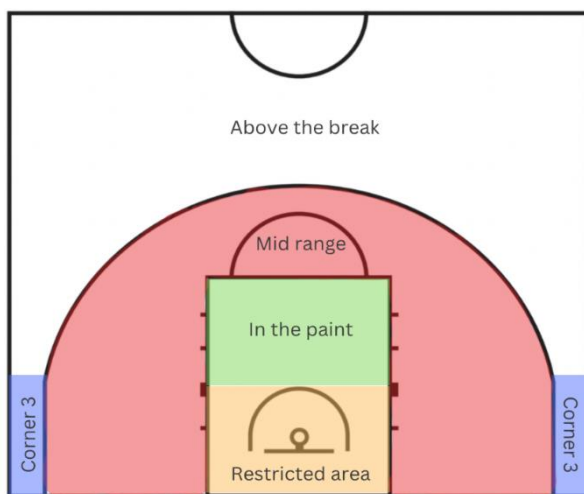


Figura 7. Esquema de les zones de llançament a cistella a la pista de bàsquet

**Above the break:** posició de llançament que anota 3 punts si s'encerta, des de les zones de 45° i centrals fora del semicercle.

**Corner 3:** posició de llançament de que anota 3 punts si s'encerta, des de les cantonades fora del semicercle.

**Mid range:** posició de llançament que anota 2 punts si s'encerta, dintre del semicercle però fora de la zona rectangular.

**In the paint:** posició de llançament que anota 2 punts si s'encerta, dintre del rectangle de la zona allunyat de la cistella.

**Restricted area:** posició de llançament que anota 2 punts si s'encerta, dintre del rectangle a prop de la cistella.

## Taula 1.

Llistat de variables seleccionades per a l'estudi de les zones de llançament a cistella.

Variable	Descripció Variable
RA_FGM	Restricted area_Field Goal Made*
RA_FGA	Restricted area_Field Goal Attempted**
RA_FG%	Restricted area_Field Goal Percentage***
IP_FGM	In the paint_Field Goal Made
IP_FGA	In the paint_Field Goal Attempted
IP_FG%	In the paint_Fiel Goal Percentage
MR_FGM	Media Range_Field Goal Made
MR_FGA	Media Range_Field Goal Attempted
MR_FG%	Media Range_Field Goal Percentage
LC_FGM	Left Corner 3_Field Goal Made
LC_FGA	Left Corner 3_Field Goal Attempted
LC_FG%	Left Corner 3_Field Goal Percentage
RC_FGM	Right Corner 3_Field Goal Made
RC_FGA	Right Corner 3_Field Goal Attempted
RC_FG%	Right Corner 3_Field Goal Percentage
CO_FGM	Corner 3_Field Goal Made
CO_FGA	Corner 3_Field Goal Attempted
CO_FG%	Corner 3_Field Goal Percentage
ATB_FGM	Above the break_Field Goal Made
ATB_FGA	Above the break_Field Goal Attempted
ATB_FG%	Above the break_Field Goal Percentage

*Nota.* Aquesta taula mostra les variables de zones de llançament a cistella. Taula d'elaboració pròpia per l'estudi.

\* **Field Goal Made:** Llançaments anotats.

\*\* **Field Goal Attempted:** Llançaments Intentats.

\*\*\* **Field Goal Percentage:** Percentatge de llançaments anotats respecte els intentats (FGM/FGA\*100).

\*\*\*\*Els valors extrets són la Mitjana per partit per a cada temporada.

Passos seguits:

- S'agafen les dades de tots els equips per a cada temporada del període estudiat. I es divideix entre *Temporada regular*<sup>6</sup> i *Playoffs*<sup>7</sup>.
- Com que la font de dades està en format *SAP*, es fa una transformació de les dades des de l'"html" a *Excel*.
- El nom de les variables el modifiquem per a que sigui més senzill de treballar i s'agrupen en una de sola, tenint en compte que en format *SAP* estan dividides en dos supra-grups.

<sup>6</sup> Tots els partits de la lliga regular de la NBA. Juguen tots els equips els uns contra els altres.

<sup>7</sup> Fase final del campionat de la NBA. Juguen en format d'eliminària directe els 8 primers classificats de cada conferència.

RESTRICTED AREA			IN THE PAINT (NON-RA)			MID-RANGE			LEFT CORNER 3.			RIGHT CORNER 3.			CORNER 3			ABOVE THE BREAK 3.		
FGM	FGA	FG%	FGM	FGA	FG%	FGM	FGA	FG%	FGM	FGA	FG%	FGM	FGA	FG%	FGM	FGA	FG%	FGM	FGA	FG%

Figura 8. Capçalera del format de presentació de les variables en SAP

En aquest pas no es modifica res de les dades i s'acaba obtenint un fitxer d'Excel ".xlsx" amb les dades brutes extretes.

TEAM	RA_FGM	RA_FGA	RA_FG%	IP_FGM	IP_FGA
Atlanta Hawks	15.9	27.7	57.5	4.0	11.2
Boston Celtics	14.6	26.1	55.9	3.1	9.1
Charlotte Hornets	14.5	25.4	57.1	4.4	11.1
Chicago Bulls	13.4	24.2	55.5	3.2	8.7
Cleveland Cavaliers	12.4	22.2	56.0	4.6	11.5

Figura 9. Exemple de les dades extretes en les 5 primeres variables creades.

- **Posició del jugador a pista**

Un altre factor que es vol comprovar si ha afectat a aquest canvi de tendència en el tir exterior és la "posició del jugador". Per tant, s'ha extret la mateixa base de dades esmentada anteriorment, però per a cada posició diferent del jugador en el bàsquet NBA ("guard", "forward" i "center").

En aquest cas, simplement s'han extret els valors de la "Temporada Regular" per a cada posició.

A aquests 3 data sets se'ls afegeix la variable "posició", per poder unificar en un de sol utilitzant el programa SAS però amb consultes d'SQL.

- **Percentatge de victòria + estadístiques tradicionals del bàsquet**

A mesura que s'avançava en l'estudi, s'afegeixen noves hipòtesis que cal contrastar. Per tant, s'extreu de la mateixa font, una base de dades amb la mateixa estructura d'equips per temporada en el període estudiat del 2001-2022 però amb un seguit de variables estadístiques que seran necessàries pels anàlisis i modelitzacions respecte el percentatge de victòria i l'evolució del llançament exterior.

A continuació, es presenta un breu resum de les variables seleccionades per l'estudi del percentatge de victòria:

**Taula 2.**

*Llistat de variables seleccionades per a l'estudi d'estadístiques del bàsquet.*

Variable	Descripció Variable
GP	Partits Jugats
W	Nº de victòries
L	Nº de derrotes
WIN%	Percentatge de victòria
PTS	Nº de Punts de Mitjana per partit
FGM	Total llançaments encertats
FGA	Total llançaments intentats
FG%	Percentatge d'encert de llançament

<b>FTM</b>	Tirs Lliures encertats
<b>FTA</b>	Tirs Lliures intentats
<b>FT%</b>	Percentatge tirs lliures
<b>OREB</b>	Rebots Ofensius
<b>DREB</b>	Rebots defensius
<b>REB</b>	Rebots
<b>AST</b>	Assistència
<b>STL</b>	Robes
<b>BLK</b>	Taps
<b>PF</b>	Faltes Personals

*Nota.* Aquesta taula mostra les variables estadístiques bàsiques necessàries per l'estudi. Els valors extrems són de mitjana per partit per a cada temporada, excepte les "victòries", "derrotes" i "percentatge de victòria" que són el Total de la temporada. Taula d'elaboració pròpia per l'estudi.

### 3.2. Depuració de les dades

En l'Excel "brut" hi ha diferents aspectes que s'han de corregir:

- Les temporades estan pestanya per pestanya (*sheet*). 21 *sheets* independents en total.
- No hi ha una variable que marqui la temporada específica.
- Com que les variables estan diferenciades per zona, falta la suma de les zones que són de 3 punts per tenir el total.
- Les variables extrems de SAP en Excel estan en format *string*<sup>8</sup> quan haurien d'estar en format numèric.
- Hi han equips que han tingut canvi de nom durant les temporades, per tant s'ha hagut de corregir el nom a anteriors temporades pel més actual. Així es poden fer les unions entre taules de forma eficient.

S'ha utilitzat SAS per importar totes les pestanyes del Excel esmentat i arreglar els 4 punts anteriors. Correcció:

- Es crea un sol *data.frame*<sup>9</sup> amb tota la informació de les diferents pestanyes.
- Per a cada pestanya, s'afegeix la variable *temporada* amb el corresponent valor (p.e. 2001\_2002). Així un cop creat un sol *data.frame* queda especificat de quina temporada és cada registre. Això permetrà poder agrupar posteriorment per aquella variable.
- Es creen les variables *T3\_FGM*, *T3\_FGA* i *T3\_FG*.

#### Taula 3.

*Variables creades a partir de les de l'extracció.*

Variable	Construcció de la variable	Descripció Variable
<b>T3_FGM</b>	CO_FGM + ATB_FGM	Suma dels llançaments encertats des de la zona <i>Corner 3</i> i <i>Above the break</i> .
<b>T3_FGA</b>	CO_FGA + ATB_FGA	Suma dels llançaments intentats des de la zona <i>Corner 3</i> + <i>Above the break</i> .
<b>T3_FG</b>	T3_FGM/T3_FGA*100	Percentatge d'encert del total de llançament de 3 punts.

*Nota.* Taula de les variables creades i la seva descripció. Taula d'elaboració pròpia per l'estudi.

<sup>8</sup> Seqüència de caràcters utilitzats per a representar un text.

<sup>9</sup> Estructura de dades amb dues dimensions en les quals es poden emmagatzemar dades de diferents tipus en columnes.



El data frame resultant “net” s’exporta en format .xlsx per poder-lo llegir en RStudio i es corregeix el quart punt:

- Es converteixen totes les variables de les zones de llançament en format numèric per poder operar amb elles.
- Es crea la variable *Temp\_num*, que obtindrà el primer valor d’any de la variable *Temporada* per poder treballar amb ella de forma numèrica.

Aquest procés de depuració es fa tant per les dades extretes per a la *Temporada regular*, com pel data set unificat de les *Posicions del jugador* i el data set amb el *percentatge de victòria* i la resta de variables estadístiques tradicionals.

### 3.3. Tècniques estadístiques

Per anar resolent les diferents hipòtesis que han anat sorgint a mesura que s’avançava el treball, s’han utilitzat les següents tècniques estadístiques:

**Taula 4.**

*Tècniques estadístiques utilitzades en l’estudi.*

<b>Tècnica estadística</b>	<b>Breu descripció</b>	<b>Test utilitzat</b>
<b>Test de Normalitat</b>	Avaluar la normalitat de les dades en una mostra menor a 50 registres.	<b>Shapiro-Wilk</b>
	Avaluar la normalitat de les dades en una mostra menor a 50 registres.	<b>Shapiro-Francia</b>
	Avaluar la normalitat de les dades en una mostra major a 50 registres.	<b>Anderson-Darling</b>
<b>Test de Variances</b>	Contrastar si dues variables es comporten amb la mateixa variabilitat.	<b>ProvaF de dues variances. Var.test en R</b>
<b>Anàlisi exploratori de les dades</b>	Visualitzar els diferents factors que poden afectar al canvi de tendència en el tir exterior.	<b>ggplot R + Visual Basic en Excel</b>
<b>Test de Mitjanes</b>	T-Student per contrastar si hi ha diferents mitjanes poblacionals en dos mostres independents.	<b>t.test d'R</b>
<b>Clusterització</b>	Agrupar els equips de la NBA en "x" grups en el que cada equip estigui en el grup on el valor mitjà de les variables de llançament de 3 pts estigui més proper.	<b>k-means d'R</b>
<b>Correlació</b>	Contrastar la correlació entre diferents variables entre si.	<b>corr.test i corrplot d'R</b>
<b>Modelització lineal</b>	La regressió implica l’ús d’una o més variables, considerades variables independents, per predir els valors d’un altre variable (la variable dependent). Les variables que estiguin molt correlacionades amb la variable dependent s’utilitzaran per predir aquesta variable.	<b>lm d'R</b>

*Nota. Tècniques estadístiques utilitzades en el treball amb la descripció i el tipus de Test utilitzat. Taula d’elaboració pròpia per l’estudi.*



## 4. ANÀLISI DESCRIPTIU DE L'EVOLUCIÓ DE LES ZONES DE LLANÇAMENT EXTERIOR

L'interrogant que es planteja el present estudi, és si realment hi ha hagut un canvi de tendència en el tir exterior com mostra la imatge de la introducció. A partir d'aquí, es vol saber quins factors han estat els més significatius per a que hagi succeït aquest canvi i si el percentatge de victòria s'ha vist condicionat pels equips que han canviat la seva tendència en el tir.

Primer, es comprova si la hipòtesi de canvi de tendència de llançament exterior es pot observar visualment. Posteriorment, es busquen quins factors es troben rellevants per després analitzar estadísticament si aquest canvi de tendència és significatiu.

### 4.1. Evolució del llançament de 3 punts per Temporada de mitjana de tots els equips NBA

#### *Hipòtesi 1*

***En el transcurs de les temporades hi ha un salt significatiu en el llançament exterior?***

S'extreu un evolutiu per Temporada dels llançaments de tirs de 3 punts intentats, estimant el valor amb la mitjana de tots els equips de la NBA.

Amb el següent gràfic, es vol buscar resposta a si hi ha una o diverses temporades que marquessin un abans i un després en l'augment del llançament exterior. S'utilitza la variable temporada amb cada valor de mitjana de llançament de 3 punts intentats (*T3\_FGA*).

Es pot comprovar que a partir de la temporada 2011/2012 hi ha un canvi de tendència en el tir de 3 punts intentat, es llança més d'aquesta posició i cada any va augmentant progressivament.

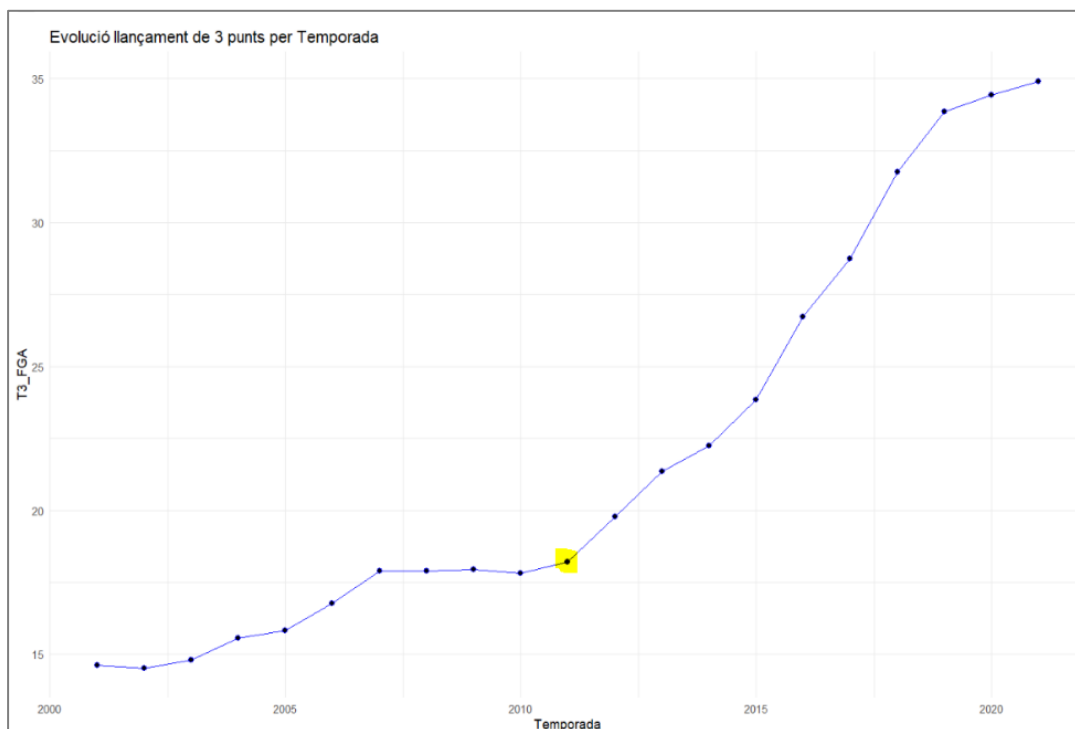


Figura 10. Canvi de tendència en el llançament de 3 PTS i % de diferència entre temporades.

## Taula 5.

*Diferències percentuals entre temporades.*

Temporada	T3_FGA	Dif % respecte temporada anterior
2001_2002	14,6	
2002_2003	14,5	-0,6%
2003_2004	14,8	1,9%
2004_2005	15,6	5,3%
2005_2006	15,8	1,6%
2006_2007	16,8	5,9%
2007_2008	17,9	6,7%
2008_2009	17,9	0,1%
2009_2010	18,0	0,3%
2010_2011	17,8	-0,7%
2011_2012	18,2	2,2%
2012_2013	19,8	<b>8,6%</b>
2013_2014	21,3	7,9%
2014_2015	22,2	4,2%
2015_2016	23,8	7,2%
2016_2017	26,7	12,1%
2017_2018	28,7	7,6%
2018_2019	31,8	10,5%
2019_2020	33,9	6,6%
2020_2021	34,4	1,7%
2021_2022	34,9	1,4%

*Nota. Diferències percentuals d'una temporada respecte la seva anterior. Taula d'elaboració pròpia per l'estudi.*

En la taula anterior s'obtenen les diferències percentuals de llançament de 3 punts (T3\_FGA) d'una temporada respecte la seva temporada anterior. S'indica el valor marcat **8.6%** com un punt d'inflexió, a partir d'aquella temporada l'augment és acumulatiu i no descendeix del 4% de creixement fins a les últimes dues temporades.

Si es divideix en dos períodes tallant per la temporada del punt d'inflexió, s'obté que l'augment del llançament de 3 punts del 2001-2012 és d'un **24.7 %** (un **2.2%** estimat per temporada de creixement) mentre que del 2012-2022 és d'un **91.67 %** (un **9.2%** estimat per temporada de creixement).

## 4.2. Evolució del percentatge d'encert de 3 punts per Temporada de mitjana de tots els equips NBA

### **Hipòtesi 2**

**Ha augmentat/disminuït el percentatge d'encert al comprovar que es llança més de 3 punts?**

Amb el següent gràfic, es vol buscar resposta a si el percentatge d'encert s'ha vist modificat degut a que es llança més d'altres zones del camp. S'utilitza la variable temporada per a cada valor de mitjana del percentatge d'encert en llançament de 3 punts (T3\_FG).

Aquest augment de tirs intentats, segons el gràfic no mostra que pugui haver afectat al percentatge d'encert, és a dir, es llança més però s'anota en mateix percentatge.

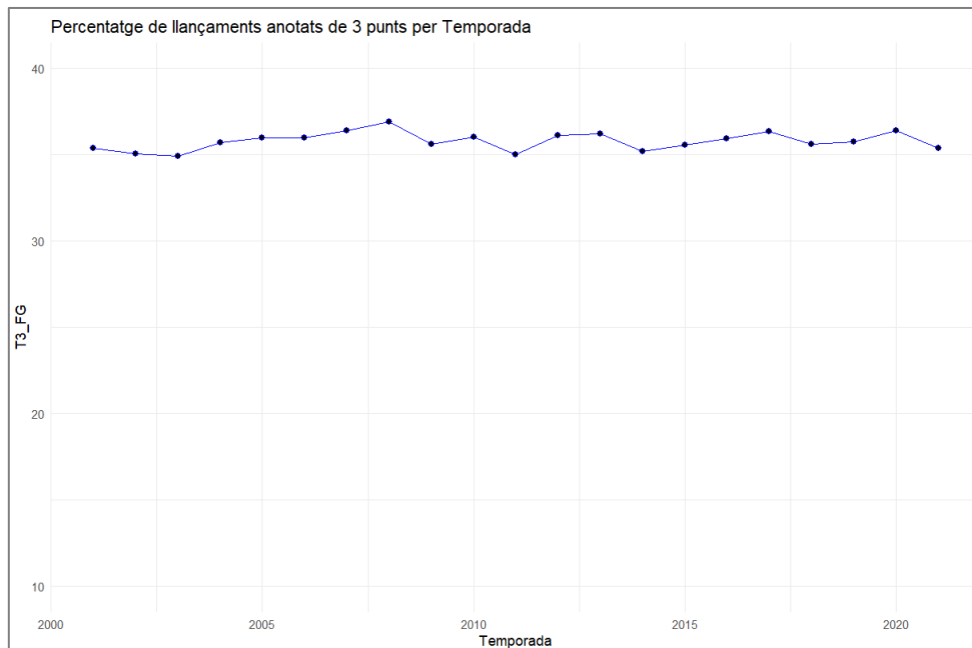


Figura 11. Percentatge llançaments anotats de 3 punts.

La diferencia percentual entre temporades no supera el  $\pm 3\%$ . Entre la primera temporada i l'última temporada només ha hagut un canvi del  $0.01\%$ .

#### 4.3. Evolució de tir intentat per les diferents zones de llançament per Temporada de mitjana de tots els equips NBA

##### **Hipòtesi 3**

***Hi ha alguna altra zona de llançament a cistella que tingui un canvi de tendència?***

Amb el següent gràfic, es vol buscar resposta a quines zones de llançament a cistella s'han vist modificades amb el pas de les temporades. S'utilitzen la variable de llançament a cistella intentats des de cada zona de llançament per temporada.

De les diferents zones de llançament, es pot observar com n'hi ha dues que tenen una variació visual per temporada.

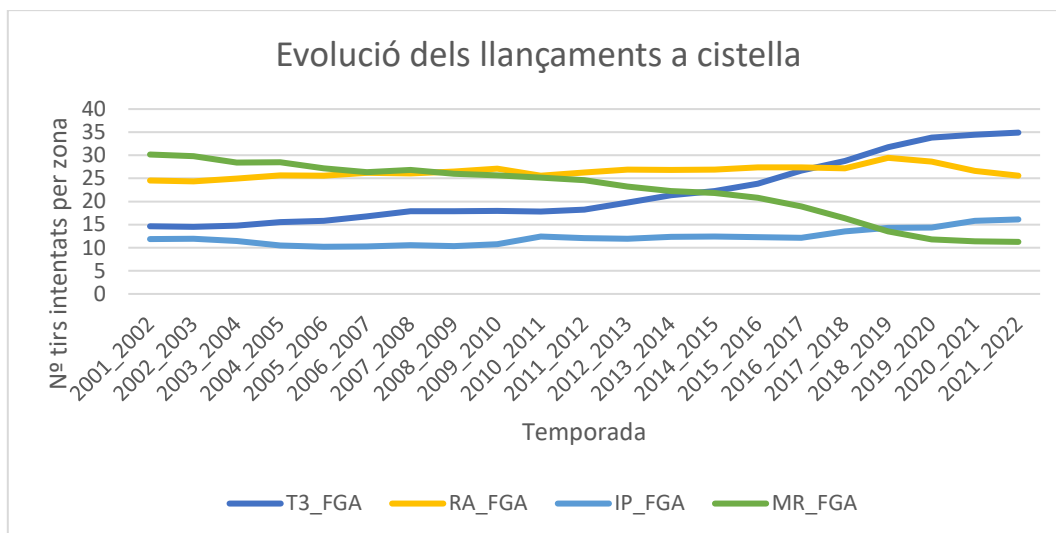


Figura 12. Tirs intentats per zona de llançament.

El nombre de tirs intentats per la zona de 3 punts ha augmentat des del 2011 com s'havia comprovat gràficament, però es confirma que el tir de mitja distància s'ha reduït. Per tant, sembla que aquest augment de tirs de 3 punts està correlacionat amb la reducció de mitja distància i no amb altres zones de llançament.

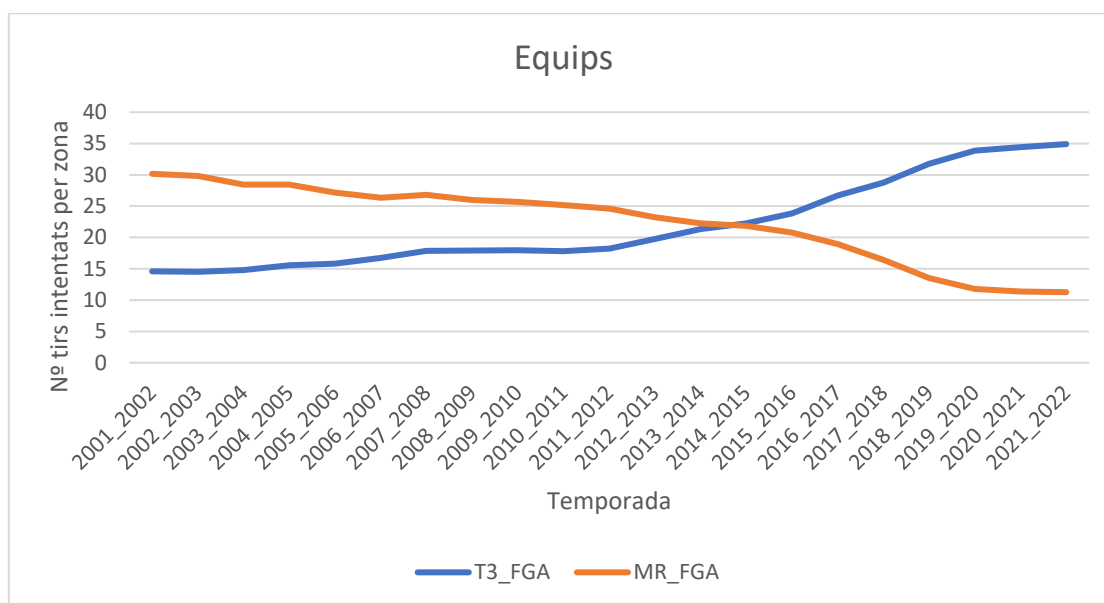


Figura 13. Zones de llançaments amb canvi de tendència significativa.

El percentatge d'encert del tir de 3 punts i el de mitja distància s'ha mantingut encara que el nombre d'intents hagi fluctuat per temporada. L'anàlisi descriptiu indica el que havia deduït Daryl Morey amb la introducció del Big Data (si es llança més de 3 punts es segueix tenint un percentatge d'encert semblant al de 2 punts i d'aquesta manera s'obté 1 punt més per cada atac), que ha contribuït a modificar la tendència dels equips a l'hora de llançar a cistella.

Es fa el mateix procediment que en l'apartat 4.1, però quedant-nos amb els percentatges del tir a mitja distància per comprovar el que s'està observant gràficament. Si es divideix en dos períodes tallant per aquella temporada, s'obté que la disminució del llançament de 3 punts del 2001-2012 és d'un **18.34 %** (un **1.6%** estimat per temporada de disminució) mentre que del 2012-2022 és d'un **51.52 %** (un **5.2%** estimat per temporada de disminució).

#### 4.4. Posició del jugador en l'evolució de llançament de 3 punts per Temporada

##### **Hipòtesi 4**

**La posició del jugador a pista ha tingut efecte en el canvi de tendència en el llançament exterior?**

Amb el següent gràfic, es vol buscar resposta a si hi ha alguna posició de jugador en concret amb una major contribució al canvi de tendència en el llançament de 3 punts.

A nivell visual no s'observa que la posició tingui especial rellevància, encara que a partir del 2011 les posicions exteriors ("guard" i "forward") comencen la tendència a l'alça com s'ha observat a nivell general, però els "center" la seva tendència a l'alça s'inicia més endavant. Com s'ha vist en els gràfics anteriors a mesura que avancen les temporades es llança més de 3 punts, així el joc ofensiu ha obert més els espais i amb el transcurs del temps els "centers" s'han vist obligats a haver de llançar també de la línia de 3 punts.

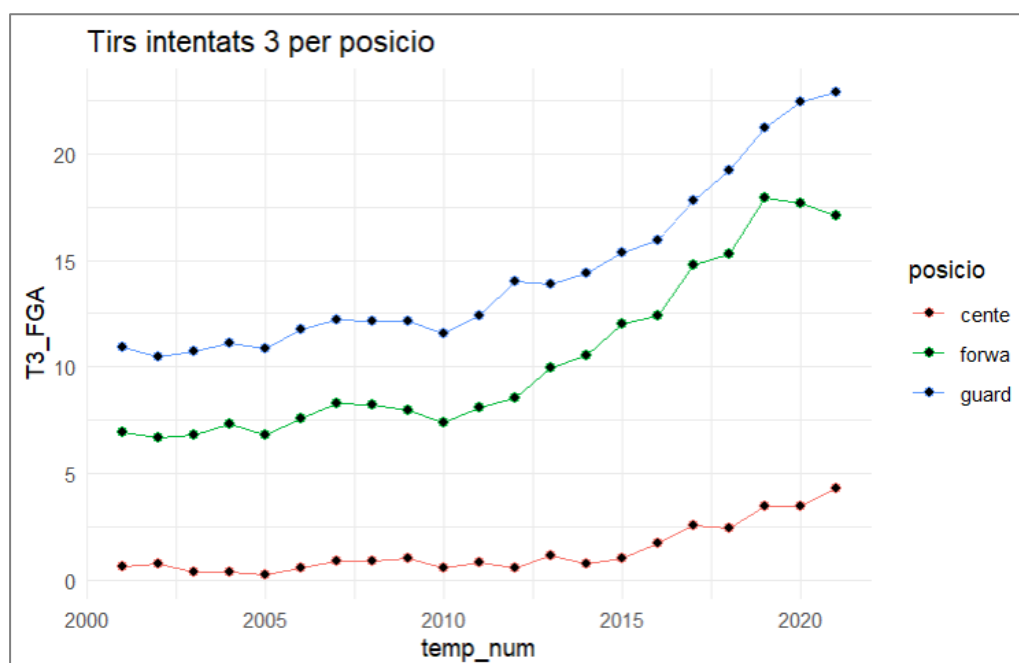


Figura 14. Tirs intentats per posició del jugador a pista.

En valors percentuals les 3 posicions augmenten a partir de la temporada 2012, com s'havia observat anteriorment, però no hi ha una posició en concret que sigui el detonant d'aquest canvi de tendència.

#### 4.5. Estudi de cas: Houston Rockets com a precursor del canvi

##### **Hipòtesi 5**

**Hi ha un equip/equips en concret que tingui més rellevància en aquest augment del llançament de 3 punts.**

Com s'ha esmentat en la Introducció i l'Estat de l'art, l'entrada de Daryl Morey a l'equip Houston Rockets va poder provocar tot aquest canvi, per aquest motiu s'estudia el cas particular d'aquest equip.

Amb el següent gràfic, es vol buscar resposta a si hi ha un canvi de tendència a l'alça en el tir de 3 a partir de l'aplicació del Big Data que va fer Morey al voltant del 2011. S'utilitzen

les mateixes variables que en el gràfic de L'Evolució de tir de tots els equips, però exclusivament pels Houston Rockets que és l'equip que va poder estar influenciat abans i amb major grau.

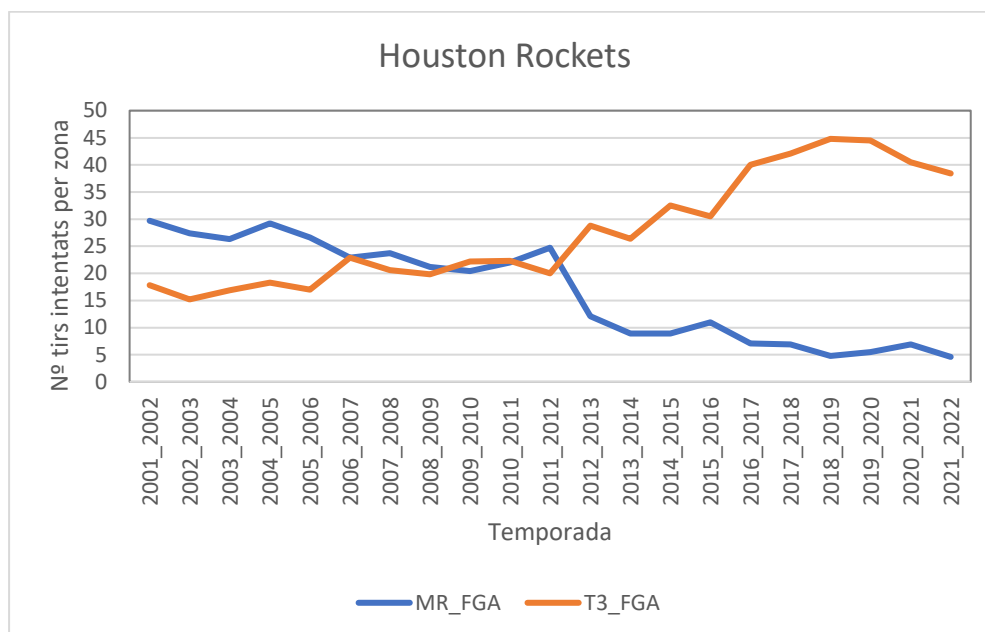


Figura 15. Zona de llançament a cistella amb canvi de tendència

Es pot observar com la tendència segueix la mitjana dels equips que s'han vist anteriorment, però amb una diferència més significativa a nivell visual. Mentre que la variable *MR\_FGA* que representa les posicions de tir de 2 punts, decreix de forma significativa a partir de la temporada 2011-2012, la *T3\_FGA* dels tirs de 3 punts creix, demostrant així el canvi de tendència del tir de 3 a l'alça donada la influència de Morey.

#### Taula 6.

*Diferències percentuals zones de llançament mitja distància vs 3 punts*

Temporada	Dif MR_FGA % respecte temporada anterior	Dif T3_FGA % respecte temporada anterior
2001_2002		
2002_2003	-8,4%	-14,6%
2003_2004	-4,2%	11,2%
2004_2005	9,9%	8,3%
2005_2006	-9,8%	-7,1%
2006_2007	-16,2%	34,7%
2007_2008	3,4%	-10,0%
2008_2009	-11,8%	-3,9%
2009_2010	-3,9%	12,1%
2010_2011	7,3%	0,5%
2011_2012	10,9%	-10,3%
2012_2013	-104,1%	44,0%
2013_2014	-36,0%	-8,3%
2014_2015	0,0%	23,1%
2015_2016	19,1%	-6,2%
2016_2017	-54,9%	31,1%

2017_2018	-2,9%	5,3%
2018_2019	-43,8%	6,4%
2019_2020	12,7%	-0,7%
2020_2021	20,3%	-9,0%
2021_2022	-50,0%	-5,2%

Nota. Diferències percentuals entre temporades per les zones de llançament que cal analitzar. Taula d'elaboració pròpia.

Si es diferencia en dos períodes tallant per la temporada 2012/2013, s'obté que l'augment del llançament de 3 punts del 2001/2012 és d'un **12.4 %** (un **1.1%** estimat per temporada de creixement) mentre que del 2012/2022 és d'un **122.5 %** (un **12.25%** estimat per temporada de creixement).

Les diferències percentuals entre temporades, en el cas de Houston Rockets, són molt més elevades que en la mitjana dels equips.

#### 4.5.1. Comparativa entre equips de la NBA en el llançament de 3 punts per Temporada.

Fent l'anàlisi per tots els equips, es pot observar que Houston Rockets (el marcat en groc) té un efecte molt significatiu en comparació amb la resta d'equips en el llançament de 3 punts. En aquesta gràfica no es pot observar si hi pot haver altres equips que també puguin haver tingut un canvi de tendència remarcable. Per tant, es planteja la clusterització per equips de la NBA utilitzant les variables de llançament de 3 punts fent servir el mètode "k-means".

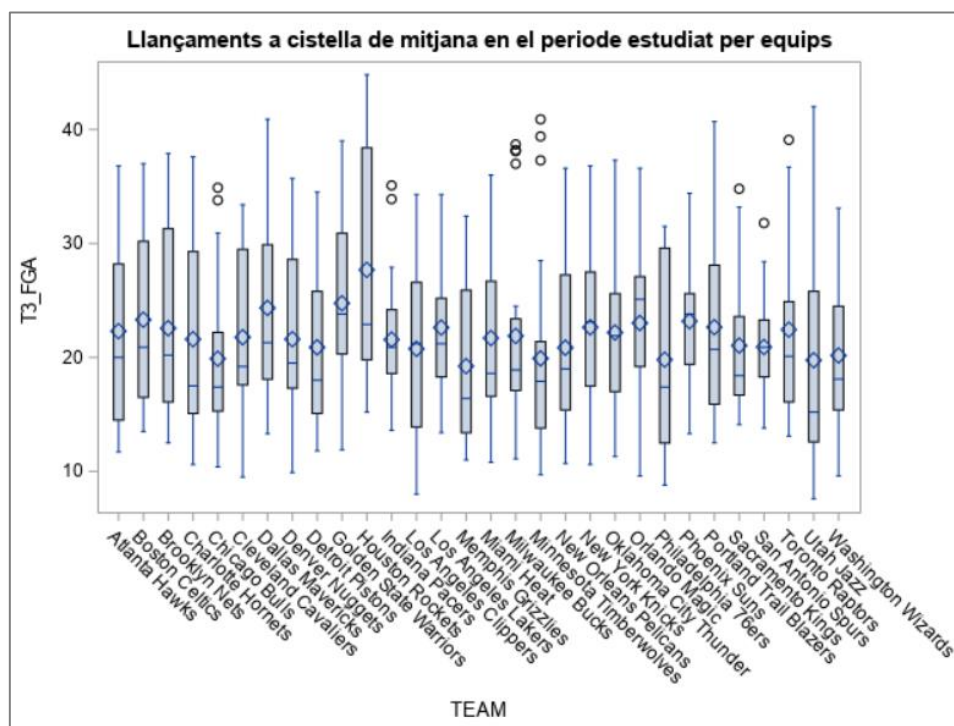


Figura 16. Llançaments a cistella de 3 punts de mitjana per les temporades del estudi per equips.

- Clusterització dels equips segons les variables de llançament de 3 punts.

El **clustering** es basa en identificar grups en les dades, de tal manera que totes les dades de cada grup (clúster) són dades de característiques similars i diferents de la resta de grups.

A continuació es presenta com s'utilitza el "*K-means*" com a sistema de clustrització. L'objectiu del "*k-means*" és agrupar observacions similars per descobrir patrons que a simple vista es desconeixen (com succeïa quan es volia observar per tots els equips la diferència en el llançament exterior).

L'algoritme intenta minimitzar la distància entre les observacions que pertanyen a un clúster i el seu centroid. És a dir, l'objectiu és minimitzar la suma de les distàncies entre els punts i el centroid que pertanyen.

Per tant, es crea un *data set* amb les variables de llançament de zona de 3 punts per a cada equip de la NBA amb la mitjana de tot el període estudiat (2001-2022).

### ***K-means*:**

1. S'inicialitzen  $k$  punts aleatoris, anomenats *centroides*.

Per determinar el  $k$  *centroides*, nombre de grups que es reparteixen els equips, s'utilitza "el mètode del colze". Consisteix en calcular i s'obté la gràfica de la suma de quadrats en cada nombre de clústers, i allà buscar un canvi de tendència per determinar el nombre òptim de clústers. En l'estudi, es decideix utilitzar el valor de  $k = 3$ .

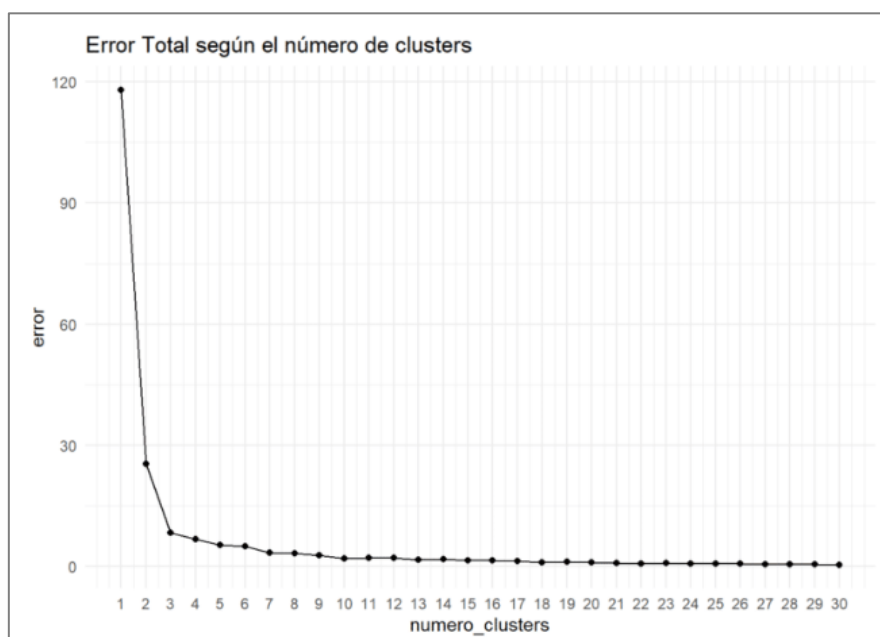


Figura 17. Taula d'error segons el nombre de clústers per determinar el nombre de grups  $k$  per clusteritzar

2. Per a cada observació, es calcula la suma d'errors al quadrat d'aquella observació respecte a cada un dels  $k$  *centroides*.
  3. Per a cada observació se l'assigna el *centroide* que menys error tingui.
- **Resultats** dels centres dels clústers i les variables utilitzades per clusteritzar:



> cl\$centers									
	CO_FGM	CO_FGA	CO_FG_	ATB_FGM	ATB_FGA	ATB_FG_	T3_FGM	T3_FGA	T3_FG
1	2.222368	5.761623	38.34035	5.782503	16.56136	34.83590	8.004872	22.32298	35.76714
2	2.018983	5.213874	38.42203	5.251753	15.08682	34.70924	7.270736	20.30069	35.65584
3	2.434921	6.225397	38.94762	6.842857	19.37460	35.61587	9.277778	25.60000	36.34603

Figura 18. Mitjana dels valors de les variables utilitzades per a cada clúster obtingut.

Els 3 grups es podrien diferenciar:

- **G2**: equips que menys llancen de 3 punts.
- **G3**: equips que més llancen de 3 punts.
- **G1**: la resta d'equips.

En G3 trobem a Houston Rockets únicament amb dos equips més.

Distribució dels equips per grup:

- **G1**: Atlanta Hawks, Boston Celtics, Charlotte Hornets, Cleveland Cavaliers, Denver Nuggets, Indiana Pacers, Los Angeles Lakers, Milwaukee Bucks, New York Knicks, Oklahoma City Thunder, Orlando Magic, Phoenix Suns, Portland Trail Blazers, Miami Heat, Toronto Raptors.
- **G2**: Chicago Bulls, Detroit Pistons, Los Angeles Clippers, Memphis Grizzlies, Minnesota Timberwolves, New Orleans Pelicans, Philadelphia 76ers, Utah Jazz, Washington Wizards, Sacramento Kings, San Antonio Spurs.
- **G3**: Dallas Mavericks, Golden State Warriors, **Houston Rockets**.

Els clústers podran servir més endavant per a posteriors anàlisis.

G3 a partir del 2012 comença l'augment significatiu en els llançaments intentats i la resta de grups comença a augmentar dos temporades posteriors, pot indicar que si que han hagut equips que han condicionat a la resta a llançar més de 3 punts.

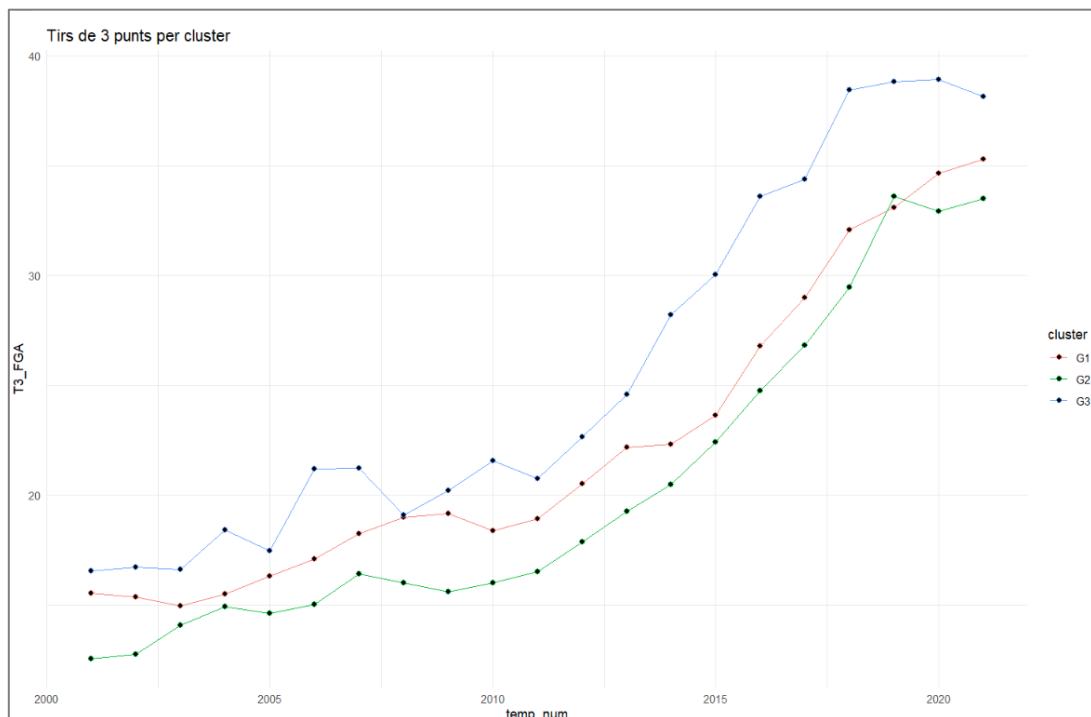


Figura 19. Llançaments intentats de 3 punts per cada clúster per temporada

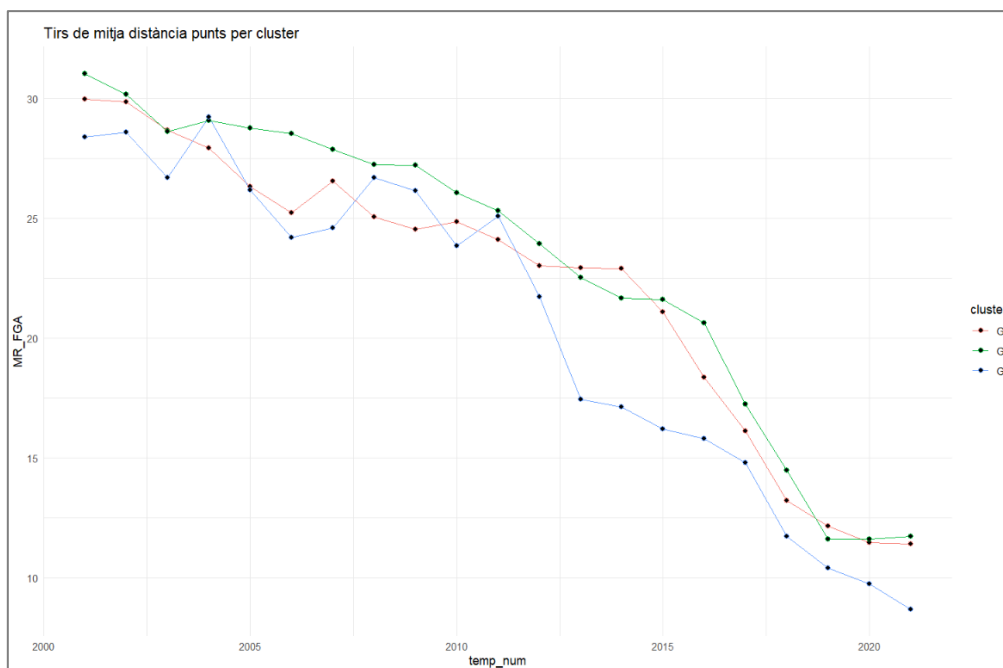


Figura 20. Llançaments intentats de mitja distància per cada clúster per temporada.

En el tir de mitjà distància, s'observa que el G3 té una baixada molt significativa a partir de l'augment del llançament de 3 punts, en canvi els equips que no llançaven tant de 3 la baixada "forta" triga unes temporades més en arribar.

Els llançaments a prop de la cistella (restricted area) no pateix aquest canvi de tendència, fluctua entre temporades. Al final el restricted area equival a fer una esmaixada o tirs molt propers a cistella que això s'ha mantingut en l'evolució del joc.

En canvi, en els llançaments de dintre de l'ampolla. Si que es pot observar que els equips G3 que més han llançat de 3 punts també han reduït els llançaments d'aquesta zona en algunes temporades i són els que menys llancen d'aquestes zones.

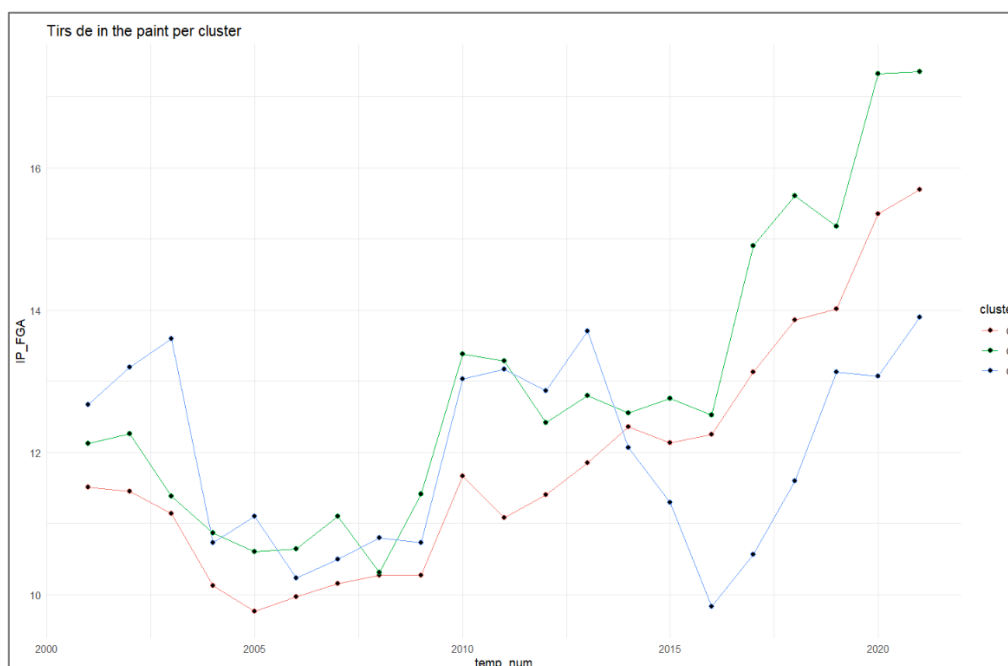


Figura 21. Llançament de dintre de l'ampolla per a cada clúster per temporada.

#### 4.6. Possibles conseqüències en el joc degut al canvi de tendència en el tir exterior

Visualment i numèricament tot fa indicar que ha hagut un augment en el llançament de 3 punts en els últims 20 anys, però que a partir de la temporada 2012/2013 és quan hi ha aquest ascens progressiu. Aquest fet fa sorgir noves hipòtesis que poden estar condicionant al joc.

##### **Hipòtesi 6:**

***Si es llança més de fora de la zona de 3 punts i es llança menys de dintre de la zona de 2 punts, ha hagut a nivell de tirs totals intentats un canvi de tendència també?***

Amb el següent gràfic, es vol buscar resposta a si en el període estudiat ha hagut un augment dels llançaments totals a cistella. Aquest augment del llançament de 3 punts pot estar contribuint a que els atacs siguin més ràpids i portin a més llançaments a cistella.

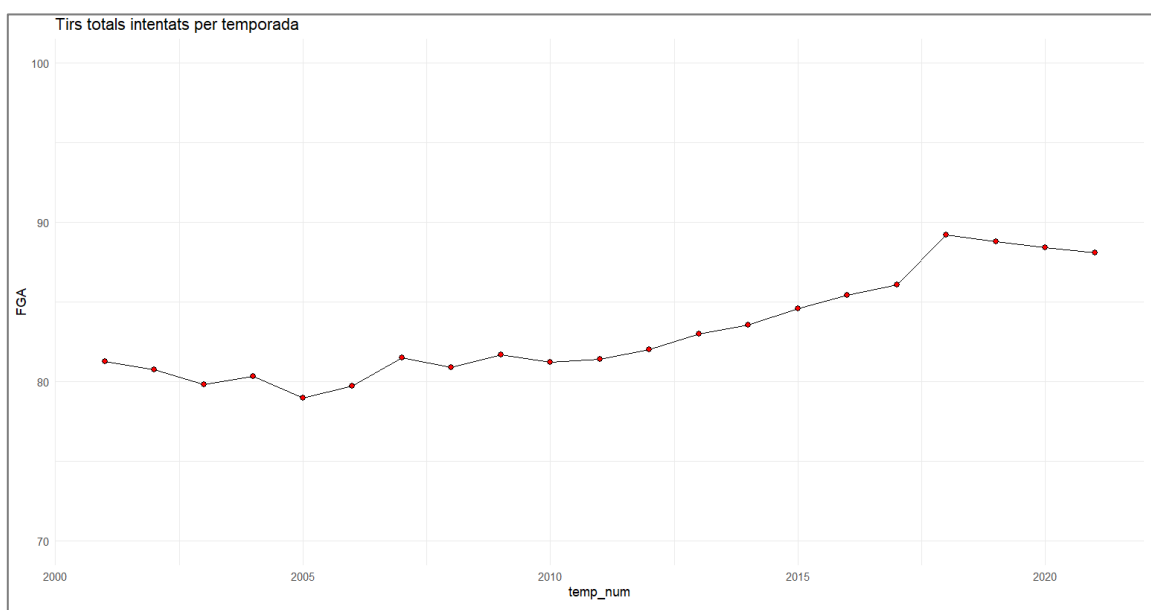


Figura 22. Llançaments totals intentats de mitjana per partit per temporada.

##### **Hipòtesi 7:**

***Com que es penetra menys a cistella, hi han menys tirs lliures tirats per equip ja que es cometen menys faltes?***

Amb el següent gràfic, es vol buscar resposta a si l'augment del llançament de 3 punts ha tingut relació a que es produeixin menys faltes de tir i d'aquesta manera menys llançaments de tir lliure. S'utilitza la variable de llançament de tir lliure intentat (*FTA*) per Temporada.

Es pot observar que hi ha una baixada menor a partir de la temporada 2011, però no segueix una tendència descendent. Es redueixen els tirs lliures, però no sembla provocat per l'augment del llançament de 3 punts o veuríem un descens més prorrogat a mesura que augmenten les temporades.

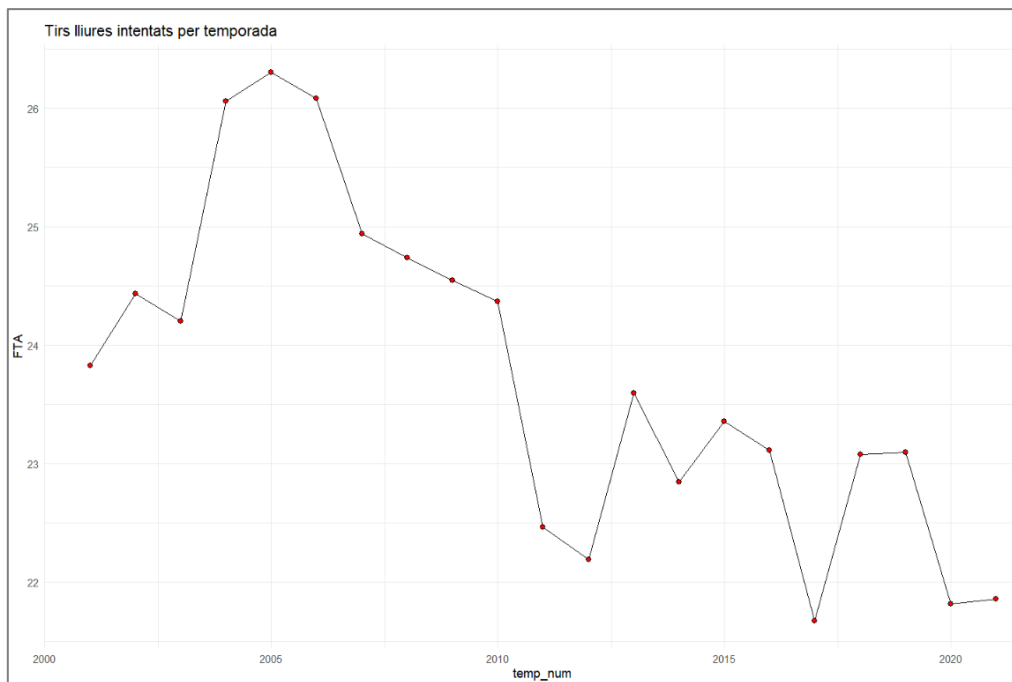


Figura 23 Llançaments de tirs lliures intentats de Mitjana.

**Hipòtesi 8:**

**Aquest augment en el llançament ha contribuït a més victòries ?**

Amb el següent gràfic, es vol buscar resposta a si l'augment de llançament a cistella de 3 punts ha portat als equips a obtenir més victòries.

En el gràfic es pot comprovar que el G3 (equips que més llancen de 3 punts) a partir de la temporada 2012 (punt d'inflexió en el creixement dels llançaments de 3 punts) han augmentat el percentatge de victòria en els següents anys de forma molt diferenciada als altres 2 grups fins la temporada 2017. En els últims anys del període estudiat, el percentatge de victòria s'ha igualat entre els 3 grups de clúster.

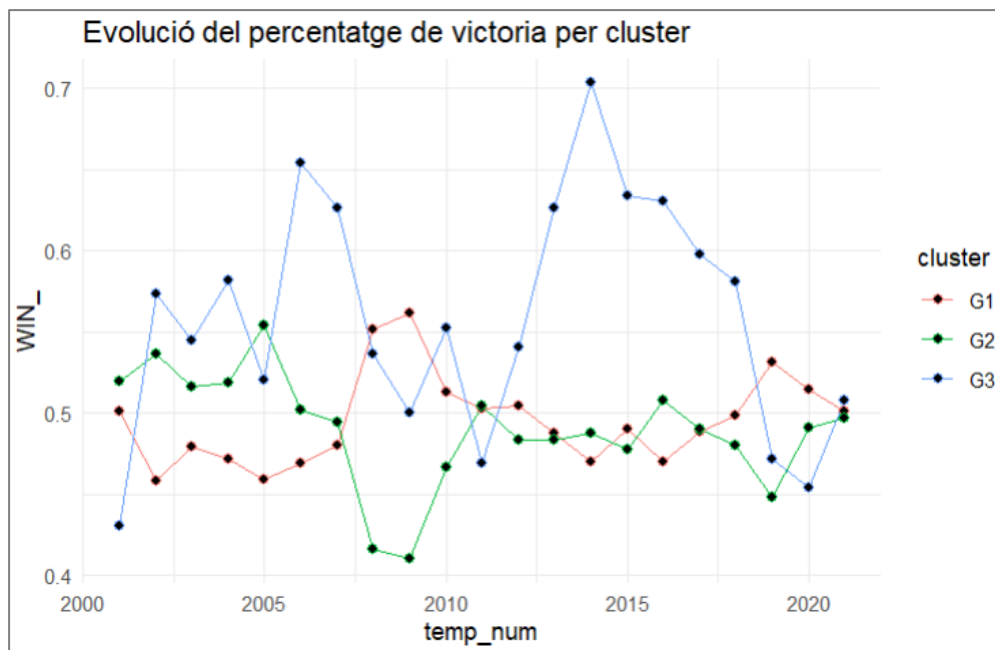


Figura 24. Percentatge de victòria per cluster.

## 5. ANÀLISI ESTADÍSTIC DE L'EVOLUCIÓ DE LES ZONES DE LLANÇAMENT EXTERIOR

### 5.1. Estudi de normalitat de la variable resposta (T3\_FGA)

En els anàlisis descriptius s'ha detectat que hi ha un canvi de tendència, al llarg de les temporades, del nombre de llançaments intentats de 3 punts, possibles factors que han pogut ser condicionants d'aquest canvi i d'altres que han pogut ser causats per aquest canvi. Per tant, primerament es fa un estudi de normalitat dels llançaments intentats de 3 punts (T3\_FGA) per determinar si les dades es comporten amb distribució normal i si es podran emprar mètodes estadístics paramètrics.

Es seleccionen 3 equips de la NBA amb diferent quocient de victòries per a una temporada en concret. Els equips són: *Golden State Warriors*, *Orlando Magic* i *Utah Jazz* per a la temporada 2017/2018, i es recullen el nombre de llançaments intentats contra cada rival d'aquella temporada per comprovar la seva distribució.

#### **Hipòtesi:**

***H<sub>0</sub>: La distribució de la variable és normal***

***H<sub>1</sub>: La distribució de la variable és diferent a la normal***

*Nivell de significança: Alpha=0.05.*

*Criteri de decisió: Si p-valor < Alpha es rebutja H<sub>0</sub>*

Per a cada equip seleccionat s'obté l'histograma de Normalitat (es pot comprovar en l'Annex). Exemple d'histograma de normalitat:

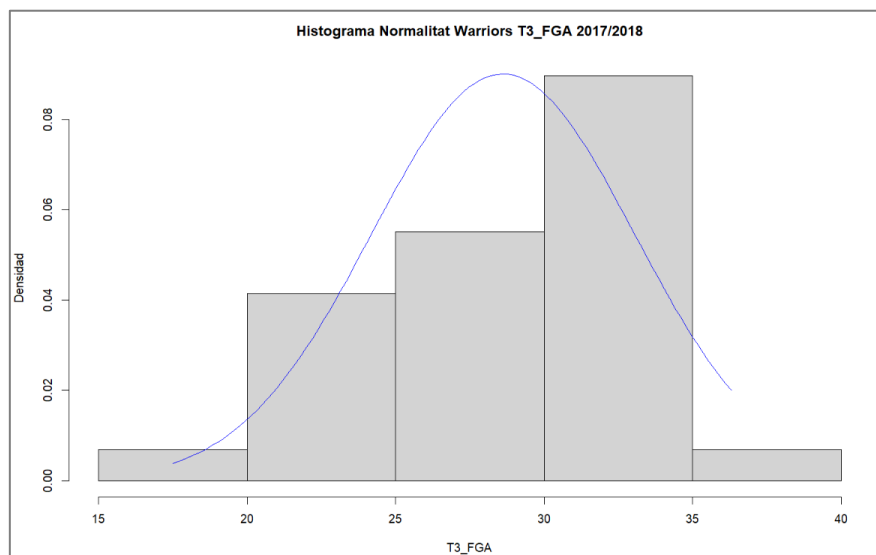


Figura 25. Histograma de normalitat dels llançaments de 3 punts intentats.

Es decideix utilitzar la prova **Shapiro-Wilk** per contrastar la normalitat del conjunt de dades. El motiu principal d'aquesta elecció és que **Shapiro-Wilk** s'utilitza quan el tamany de la mostra a analitzar és reduït ( $\leq 50$  registres) com en el cas d'aquest estudi.

Exemple de sortida en R del test **Shapiro-Wilk**:

```

Shapiro-Wilk normality test
data: normalitat[normalitat$equip == "warriors" & normalitat$Temporada == "2017_2018", "3PA"]
W = 0.96492, p-value = 0.4315

```

Figura 26. Resultats test de Normalitat en R.

Taula de valors de la prova **Shapiro-Wilk** per a cada contrast:

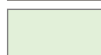
### Taula 7.

Resultats test de Normalitat llançament de 3 punts.

Temporada	Equip	Variable	W	p-valor
2017_2018	Golden State Warriors	T3_FGA	0,96	0,43
2017_2018	Orlando Magic	T3_FGA	0,93	0,21
2017_2018	Utah Jazz	T3_FGA	0,95	0,18



Ombreiat taronja: No es pot rebutiar H0. la distribució de la variable és Normal.



Ombreiat verd: Es rebutia H0. la distribució de la variable no és Normal.

Nota. Esquema resultats test de Normalitat pels equips seleccionats. Taula d'elaboració pròpia per l'estudi.

Els resultats de la prova són tots amb  $p\text{-valor} > \text{Alpha}$ , per tant s'accepta  $H_0$  i es pot confirmar que la variable T3\_FGA prové d'una distribució normal.

- **Estudi de normalitat de la resta de variables significatives en l'estudi:**

El mateix estudi de normalitat esmentat anteriorment, s'ha fet per a la resta de variables que s'han comprovat en l'anàlisi descriptiu que poden ser considerades factors decisius en el canvi de tendència agafant únicament un dels equips seleccionats i per a tot el període.

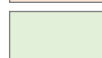
### Taula 8.

Resultats test de Normalitat de les diferents variables seleccionades.

Temporades_test	Equip	Variable	W	p-valor
2017-2018	Golden State Warriors	T3_FGA	0,96	0,43
	Golden State Warriors	T3_FG	0,97	0,72
	Golden State Warriors	MR_FGA	0,92	0,1
	Golden State Warriors	IP_FGA	0,95	0,3
	Golden State Warriors	RA_FGA	0,90	0,08
	Golden State Warriors	FGA	0,99	0,95
	Golden State Warriors	FTA	0,96	0,41
	Golden State Warriors	WIN_	0,97	0,73



Ombreiat taronja: No es pot rebutiar H0. la distribució de la variable és Normal.



Ombreiat verd: Es rebutia H0. la distribució de la variable no és Normal.

Nota. Esquema resultats test de normalitat de l'equip seleccionat amb les variables significatives. Taula d'elaboració pròpia per l'estudi.

Totes les variables tenen un  $p\text{-valor} > \text{Alpha}$ , per tant s'accepta  $H_0$  i es pot confirmar que totes les variables provenen d'una distribució normal.

## 5.2. Test de mitjanes de la variable "Temporada" per comprovar la seva significativitat

S'ha pogut comprovar visualment que hi ha hagut un augment en el llançament de 3 punts, però es vol contrastar que aquest canvi és estadísticament significatiu. Per tant, es compara mitjançant un test de mitjanes entre la primera temporada (2001/2002) contra la última temporada (2021/2022).

El contrast que es planteja per a comparar les mitjanes de dos poblacions és:

$$\begin{cases} H_0 : \mu_1 - \mu_2 = 0 \\ H_1 : \mu_1 - \mu_2 \neq 0 \end{cases}$$

Per poder aplicar aquest tipus de test, primer s'ha de comprovar si es compleix la normalitat i la homogeneïtat de les variàncies de les mostres poblacionals a l'estudi.

- **Normalitat de les dades:**

Aplicant Shapiro-Wilk com anteriorment s'ha esmentat per a les dues temporades, s'accepta  $H_0$  i es determina que tenen una distribució normal.

- **Hipòtesi d'homogeneïtat de les variàncies:**

$$\begin{cases} H_0 : \sigma_1 = \sigma_2 \\ H_1 : \sigma_1 \neq \sigma_2 \end{cases}$$

Nivell de significança:  $\text{Alpha}=0.10$ .

Criteri de decisió: Si  $p\text{-valor} < \text{Alpha}$  es rebutja  $H_0$

S'utilitza un contrast d'igualtat de variàncies poblacionals aplicant un *var.test d'R* en les dades a analitzar.

```
> var.test(test_mitjanes[, "T3_FGA"] ~ test_mitjanes[, "Temporada"])

      F test to compare two variances

data:  test_mitjanes[, "T3_FGA"] by test_mitjanes[, "Temporada"]
F = 1.329, num df = 28, denom df = 29, p-value = 0.4508
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.6297995 2.8184972
sample estimates:
ratio of variances
      1.329023
```

Figura 27 Resultats test de Variàncies entre dues variables.

El  $p\text{-valor}$  és de 0.45, per tant no es pot rebutjar  $H_0$  i s'assumeix que les variàncies són iguals per ambdós grups.

- **Test t de Student:**

Com s'ha verificat en els apartats anteriors que es compleixen les hipòtesis de normalitat i homoscedasticitat, es pot aplicar el test t de Student utilitzant el *t.test d'R*.

```

Two Sample t-test

data: test_mitjanes[, "T3_FGA"] by test_mitjanes[, "Temporada"]
t = -24.026, df = 57, p-value < 2.2e-16
alternative hypothesis: true difference in means between group 2001_2002 and group 2021_2022 is not equal to 0
95 percent confidence interval:
 -21.99153 -18.60778
sample estimates:
mean in group 2001_2002 mean in group 2021_2022
      14.61034             34.91000

```

Figura 28 Resultats test de mitjanes entre T3\_FGA vs Temporada.

El *p*-valor és de **2.2e-16**, és pot rebutjar  $H_0$  i determinar que les mitjanes són diferents entre ambdós grups.

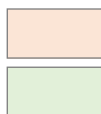
- **Test de mitjanes de la resta de variables significatives en l'estudi**

El mateix test es fa per a la resta de zones de llançament i variables que són importants per resoldre les hipòtesis que s'han plantejat en la part d'anàlisi descriptiu, on es pot determinar que les zones de llançament de *mitja distancia* (MR\_FGA), *dintre de la pintura* (IP\_FGA), els llançaments totals intentats (FGA) i els llançaments de tirs lliures intentats (FTA) es rebutja  $H_0$  i les mitjanes són diferents entre ambdós grups, però en canvi els llançaments de sota de la cistella (RA\_FGA) i el percentatge de llançament de 3 punts (T3\_FG) no es pot rebutjar  $H_0$  i les mitjanes són iguals en ambdós grups.

**Taula 9.**

*Test de Mitjanes de les variables rellevants de l'estudi.*

Temporades_test	Variable	t	p-valor
2001/2002 vs 2021/2022	MR_FGA	22	2,20E-16
	IP_FGA	-5,99	1,49E-07
	RA_FGA	-1,47	0,15
	T3_FG	-0,38	0,71
	FGA	-10,82	1,92E-15
	FTA	3,88	0,0003



Ombreiat taronja: No es pot rebutjar  $H_0$ . la distribució de la variable és Normal.

Ombreiat verd: Es rebutia  $H_0$ . la distribució de la variable no és Normal.

*Nota. Resultats test de mitjanes de les variables rellevants d'estudi. Taula d'elaboració pròpia per l'estudi.*

**5.3. Correlació entre el llançament de 3 punts vs mitja distancia**

En l'anàlisi descriptiu s'ha pogut comprovar que l'augment del llançament de 3 punts (T3\_FGA) sembla que ha fet reduir el llançament de mitja distancia (MR\_FGA), per tant sembla que les dues variables estan correlacionades entre si. Per tant, es calcula la correlació entre elles i s'utilitza un test per analitzar la significança de la pròpia correlació. S'utilitza el *cor.test d'R* que utilitza un contrast *t-student* a partir del valor de la correlació de Pearson d'ambdós variables.



```

Pearson's product-moment correlation
data: comp_total_any[, "T3_FGA"] and comp_total_any[, "MR_FGA"]
t = -41.357, df = 19, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
-0.9978098 -0.9861810
sample estimates:
cor
-0.9944915

```

Figura 29. Resultat del test de correlació entre T3\_FGA i MR\_FGA.

El valor de la correlació entre les dues variables és  $\approx -1$  amb un *p*-valor **2.2e-16**, per tant s'accepta que hi ha correlació negativa entre les dues variables. Quan el llançament de 3 punts (T3\_FGA) augmenta disminueix el llançament de mitja distancia (MR\_FGA).

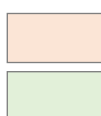
- **Correlació entre les variables més significatives**

Per poder donar resposta a les diferents hipòtesis que han anat sorgint és necessari també comprovar les correlacions entre les diferents variables, d'aquesta manera servirà també a l'hora d'introduir-les en la modelització lineal posterior.

**Taula 10.**

*Resultats correlacions respecte el llançament de 3 punts*

Variable		Valor correlació	t	p-valor
T3_FGA	IP_FGA	0,89	8,35	<b>8,81E-08</b>
T3_FGA	RA_FGA	0,64	3,63	<b>0,002</b>
T3_FGA	T3_FG	0,15	0,68	<b>0,51</b>
T3_FGA	WIN_	0,12	2,91	<b>0,003</b>
T3_FGA	REB	0,49	14,12	<b>2,20E-16</b>
T3_FGA	FGA	0,75	28,46	<b>2,20E-16</b>
T3_FGA	FTA	-0,34	-9,09	<b>2,20E-16</b>



Ombreiat taronja: No es pot rebutjar H0, la distribució de la variable és Normal.

Ombreiat verd: Es rebutja H0, la distribució de la variable no és Normal.

*Nota. Resultats dels test de correlacions entre les variables rellevants de l'estudi respecte el llançament de 3 punts. Taula d'elaboració pròpia per l'estudi.*

#### 5.4. Model lineal amb variable resposta T3\_FGA

Per acabar de confirmar quins factors han determinat més el canvi de tendència en el llançament de 3 punts, es decideix aplicar *regressió lineal* per tal de predir els valors de la **variable resposta** (en aquest cas T3\_FGA) utilitzant una o més **variables predictores**. Les variables que estan fortament correlacionades amb la variable dependent s'utilitzaran per predir aquesta variable.

Per afegir més informació a les ja correlacions calculades en l'apartat anterior, s'utilitza la funció *corrplot d'R* per elaborar un **diagrama de correlació** per acabar d'explorar les dades i confirmar quines variables tenen més interacció amb la variable resposta.

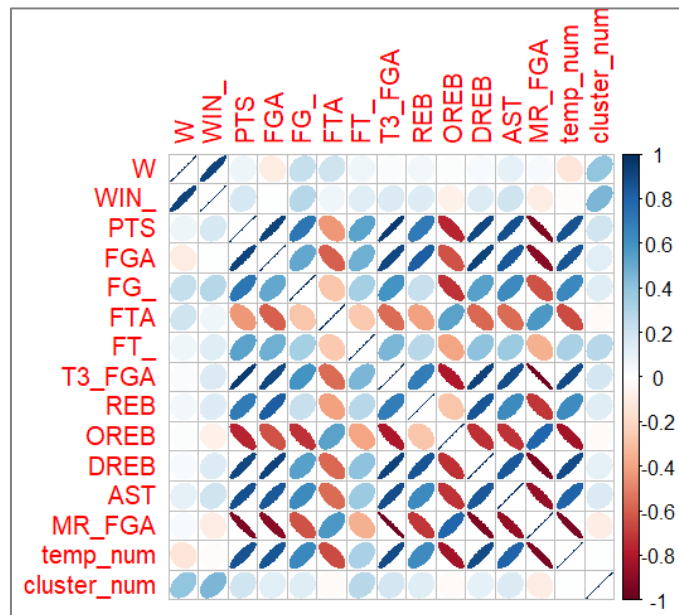


Figura 30. Mapa de correlacions entre la variable T3\_FGA amb la resta de significatives.

Un cop s'han fet les diferents proves amb diferents modelitzacions utilitzant totes les variables que s'han anat veient en l'estudi, s'ha acabat obtenint el millor model possible per predir els llançaments intentats de 3 punts (*T3\_FGA*), que és el següent:

$y = T3\_FGA$  (numèrica).  
 $X_1 = \text{clúster\_num}$  (numèrica),  $X_2 = MR\_FGA$  (numèrica),  $X_3 = IP\_FGA$  (numèrica),  $X_4 = RA\_FGA$  (numèrica).

```
Call:
lm(formula = T3_FGA ~ cluster_num + MR_FGA + IP_FGA + RA_FGA,
    data = comp_total_cluster)

Residuals:
    Min       1Q   Median       3Q      Max
-3.05860 -0.71275  0.07202  0.64947  2.29935

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  74.12357    4.29044   17.276 < 2e-16 ***
cluster_num   0.37426    0.19412    1.928  0.0588 .
MR_FGA      -1.31126    0.03348  -39.161 < 2e-16 ***
IP_FGA      -0.53880    0.10750   -5.012  5.38e-06 ***
RA_FGA      -0.63663    0.12362   -5.150  3.26e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.166 on 58 degrees of freedom
Multiple R-squared:  0.978,    Adjusted R-squared:  0.9765
F-statistic: 645.1 on 4 and 58 DF,  p-value: < 2.2e-16
```

Figura 31. Resultat de la modelització lineal amb variable resposta llançament de 3 punts intentats.

S'obté un coeficient de determinació molt proper a 1, per tant el model ajusta de forma molt eficient els llançaments a cistella de 3 intentats.

Tant el model com les variables predictores tenen *p-valors* amb significació menor al valor d'Alpha i per tant significatives en el model.

El model compleix els supòsits dels residus que ha de tenir. (Veure en l'Annex).

1. La relació entre les variables X e Y deu ser aproximadament lineal.
2. Els residus es distribueixen segons una  $N(0, \sigma^2)$ .
3. Els residus són independents uns dels altres.
4. Els residus tenen variança constant.

El **coeficient de determinació (R2)** determina quin percentatge de la variació total en les variables independents pot ser explicada per la variació en la variable dependent. En aquest cas, el **coeficient de determinació (R2)** és de 0.978, per tant un 2.2% de les variables aplicades no estan relacionades amb la variable T3\_FGA.

Els resultats obtinguts indiquen quins són els factors que han pogut ser causants o derivats del canvi en el llançament de 3 punts.

\*Els diferents models lineals provats amb les diferents variables que s'han anat veient en l'estudi, es poden trobar en l'Annex.

### 5.5. Model lineal amb variable resposta WIN\_

Per finalitzar, s'acaba de determinar si aquest canvi contribueix a un canvi en el percentatge de victòria, o bé quins factors són més determinants per augmentar el percentatge de victòria. Es decideix aplicar *regressió lineal* per tal de predir els valors de la **variable resposta** (en aquest cas WIN\_) utilitzant una o més **variables predictores**. Les variables que estan fortament correlacionades amb la variable dependent s'utilitzaran per predir aquesta variable.

Per afegir més informació a les ja correlacions calculades en l'apartat anterior, s'utilitza la funció *corrplot d'R* per elaborar un **diagrama de correlació** per acabar d'explorar les dades i confirmar quines variables tenen més interacció amb la variable resposta.

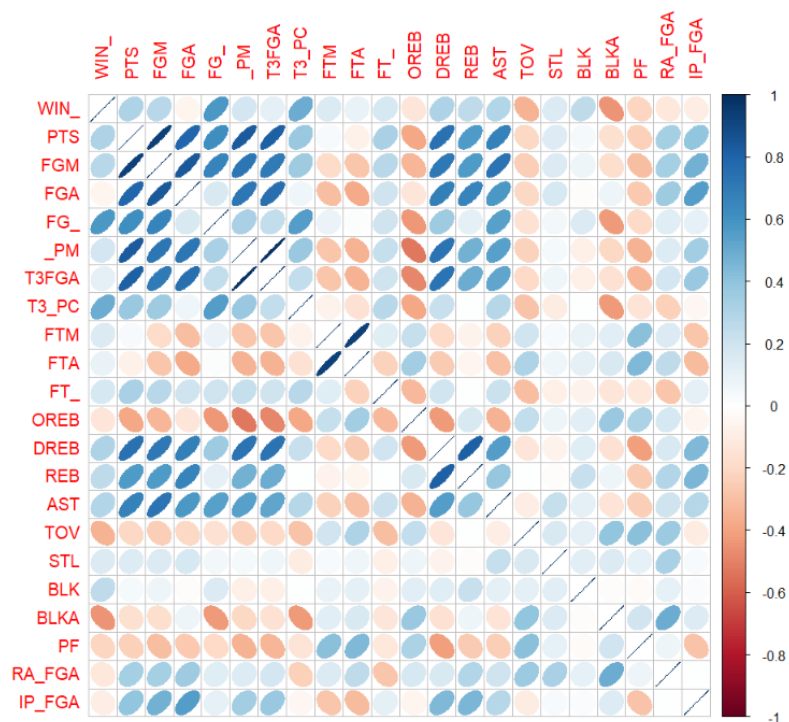


Figura 32. Mapa de correlacions entre el percentatge de victòria i la resta de variables.

Un cop s'han fet les diferents proves amb diferents modelitzacions utilitzant totes les variables que s'han anat veient en l'estudi, s'ha acabat obtenint el millor model possible per predir el percentatge de victòria (*WIN\_*), que és el següent:

$y = \text{WIN}_$  (numèrica).  
 $X_1 = \text{T3\_PC}$  (numèrica),  $X_2 = \text{FGA}$  (numèrica),  $X_3 = \text{PTS}$  (numèrica),  $X_4 = \text{REB}$  (numèrica),  $X_5 = \text{AST}$  (numèrica),  $X_6 = \text{BLK}$  (numèrica),  $X_7 = \text{TOV}$  (numèrica),  $X_8 = \text{STL}$  (numèrica),  $X_9 = \text{STL}$  (numèrica),  $X_{10} = \text{FTA}$  (numèrica),  $X_{11} = \text{T3\_FGA}$ .

```
Call:
lm(formula = WIN_ ~ T3_FGA + T3_PC + FGA + PTS + REB + AST +
    BLK + TOV + STL + FTA, data = definitiu)

Residuals:
    Min       1Q   Median       3Q      Max
-0.183625 -0.043564 -0.000998  0.044228  0.201441

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.931749   0.133001    7.006 6.47e-12 ***
T3_FGA      -0.002835   0.000719   -3.943 8.98e-05 ***
T3_PC       0.011834   0.001914    6.181 1.16e-09 ***
FGA        -0.053051   0.001781  -29.787 < 2e-16 ***
PTS         0.018097   0.001324   13.672 < 2e-16 ***
REB         0.051821   0.001897   27.318 < 2e-16 ***
AST         0.007081   0.001835    3.858 0.000126 ***
BLK         0.007140   0.003573    1.999 0.046101 *
TOV        -0.058032   0.002555  -22.712 < 2e-16 ***
STL         0.067005   0.003422   19.579 < 2e-16 ***
FTA        -0.012422   0.001731   -7.176 2.07e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.06432 on 616 degrees of freedom
Multiple R-squared:  0.8174,    Adjusted R-squared:  0.8144
F-statistic: 275.7 on 10 and 616 DF,  p-value: < 2.2e-16
```

Figura 33. Resultat de la modelització lineal amb variable resposta percentatge de victòria.

S'obté un coeficient de determinació proper a 1, per tant el model ajusta eficient el percentatge de victòria.

Tant el model com les variables predictores tenen *p-valors* amb significació menor al valor d'Alpha i per tant significatives en el model.

El model compleix els supòsits dels residus que ha de tenir.

El **coeficient de determinació** (*R*<sup>2</sup>) determina quin percentatge de la variació total en les variables independents pot ser explicada per la variació en la variable dependent. En aquest cas, el **coeficient de determinació** (*R*<sup>2</sup>) és de 0.84, per tant un 16% de les variables aplicades no estan relacionades amb la variable *T3\_FGA*.

## 6. DISCUSSIÓ DE RESULTATS I CONCLUSIONS

En l'elaboració d'aquest TFG s'ha partit d'una imatge, extreta d'una font periodística, de llançaments a cistella on s'observava un canvi de tendència en el tir del darrer any respecte al primer del segle XXI, aquest fet ha portat al plantejament d'una hipòtesi inicial: Esbrinar si l'estudi i la implementació del Big Data que estaven practicant els "Houston Rockets" i d'altres equips havien portat a aquesta modificació del joc i si realment existia un canvi de tendència significatiu.

Durant el treball de recopilació i anàlisi han anat sorgint d'altres hipòtesis relacionades amb la hipòtesi inicial, sobre les quals, amb els resultats obtinguts, s'ha avaluat si han pogut ser corroborades:

*Hipòtesi 1: En el transcurs de les temporades hi ha un salt significatiu en el llançament exterior de 3 punts?*

A partir dels resultats obtinguts, es pot afirmar que hi ha hagut un salt significatiu en el llançament a cistella del darrer any en comparació entre el primer del període estudiat (2001 vs 2022). També es pot determinar que a partir del 2012 el creixement en el llançament exterior augmenta percentualment de forma més accentuada comparat amb els anys anteriors, seguint una progressió aritmètica.

Altrament, el transcurs de les temporades no han sigut un factor decisiu en el canvi d'aquesta tendència veient la modelització resultant.

*Hipòtesi 2: Ha augmentat/disminuït el percentatge d'encert al comprovar que es llança més de 3 punts?*

S'ha corroborat que el percentatge d'encert s'ha mantingut en uns mateixos valors, per tant, el fet d'haver llançat més a cistella no ha provocat que s'encistelli més o menys. Com que els percentatges s'han mantingut i el llançament de 3 punts val 1 punt més que el de mitja distància, això ha contribuït a que els equips hagin pogut apostar més pel tir de més enllà de la zona.

*Hipòtesi 3: Hi ha alguna altra zona de llançament a cistella que tingui un canvi de tendència?*

L'augment en el llançament de 3 punts és inversament proporcional al llançament de mitja distància, per tant el llançament de mitja distància té un canvi de tendència negatiu en el transcurs de les temporades.

*Hipòtesi 4: La posició del jugador a pista ha tingut efecte en el canvi de tendència en el llançament exterior?*

S'ha pogut determinar que la posició del jugador no ha sigut un factor determinant en el canvi de tendència en el llançament de 3 punts, és a dir, no hi ha una posició en concret que hagi provocat l'augment de les altres posicions o bé una sola posició que acapari tot l'augment del llançament exterior.

*Hipòtesi 5: Hi ha un equip/equips en concret que tingui més rellevància en aquest augment del llançament de 3 punts.*

Els resultats obtinguts ens indiquen que si que hi han equips que han tingut més rellevància en aquest canvi de tendència i que han sigut un factor determinant d'aquest augment. Sobretot l'equip analitzat en particular Houston Rockets (que formava part del clúster dels que més llancen), és l'equip amb més rellevància a l'hora de llançar més de 3 punts que de mitja distància.

Hipòtesi 6: *Si es llança més de fora de la zona de 3 punts i es llança menys de dintre de la zona de 2 punts, ha hagut a nivell de tirs totals intentats un canvi de tendència també?*

Es llança quelcom més a cistella amb el transcurs de les temporades, però no es determina significativament que hi hagi un augment dels tirs totals intentats. Per tant, tot fa indicar que l'augment del llançament de 3 punts va lligat al descens del llançament de mitja distància. Aquest augment no és significatiu. Pot ser degut a que l'anàlisi s'ha fet a partir del valor mitjà de tots els equips i s'equilibra molt més la fluctuació en el llançament. Per exemple, observant el cas particular dels Houston Rockets, si que s'observa que es llança més a cistella a nivell total, però com s'ha esmentat a l'Estat de l'art, un dels seus objectius eren atacs més ràpids, però en contraposició hi ha d'altres equips que s'han mantingut amb el mateix nombre de llançaments a cistella per partit i això fa que no es pugui corroborar la significativitat de l'augment de llançament totals.

Hipòtesi 7: *Com que es penetra menys a cistella, hi han menys tirs lliures tirats per equip ja que es cometen menys faltes?*

Els valors de correlació fan indicar que com més llançaments de 3 punts hi ha en un partit per un equip, aquest equip llança menys tirs lliures. Per tant, com que hi ha hagut aquest canvi de tendència en el llançament a cistella en el transcurs de les temporades, també hi ha hagut un descens en el llançament de tirs lliures, però poc significatiu. Cal afegir, que ha faltat acabar de corroborar-ho introduint altres variables al model predictiu que podrien estar afectant, com ara la fluctuació del nombre de faltes personals de cada equip o les zones més "pitades" d'aquestes faltes, tot per falta de dades que es podrien arribar a obtenir per a posteriors estudis sobre el tema.

Hipòtesi 8: *Aquest augment en el llançament ha contribuït a més victòries?*

D'inici, observant el model resultant, l'augment en el nombre de llançaments intentats no contribueix a més victòries. Sorprenentment, en els resultats es pot determinar que l'augment en el llançament de 3 disminueix el percentatge de victòria, però això és degut a que el propi model ens confirma que com més elevat és el percentatge d'encert d'aquests llançaments si que augmenta el percentatge de victòria en major escala. És a dir, com que si augmenta el llançament de 3 punts s'ha comprovat que el percentatge de llançament es manté i aquesta variable augmenta més el percentatge de victòria que el que fa disminuir, el que indica el model és que llançant més de 3 punts podem augmentar les victòries.

En l'agrupació dels clústers si que es podia comprovar que el grup que llançava més de 3 punts, havia augmentat el percentatge de victòria que al temps augmentaven els llançaments intentats de 3 punts. Per tant, aquests equips van introduir aquesta forma de fer i van començar a assolir més victòries, la resta va començar a augmentar també el llançament de 3 punts i el percentatge entre tots es va igualar, d'aquí que l'estudi amb tots els equips no surti significatiu. En posteriors anàlisis, s'hauria d'analitzar per a cada equip, de forma detallada, obtenint tots els resultats de totes les jornades de lliga per a cada temporada, així s'obtidria una base de dades més gran per poder corroborar de forma més exacta aquesta hipòtesi.

En conclusió, amb l'estudi realitzat es pot afirmar que ha hagut un canvi de tendència en el llançament a cistella durant el Segle XXI. Els equips, sobretot al llarg de l'última dècada, han passat a llançar molt més de 3 punts i reduir de forma proporcional els llançaments de mitja distància. Com que els percentatges no es veien alterats, els equips obtenen un punt més per a cada llançament, aportant així, més possibilitats de victòria. Aquesta condició ha fet que els llançaments de dintre de l'ampolla també s'hagin vist condicionats ja que si es penetra a cistella és per acabar amb finalització propera o llançament de dintre de la pintura. Aquest canvi de forma de jugar, s'ha vist que ha condicionat a totes les posicions dels jugadors al camp, on els propis pivots també intenten llançaments exteriors.

El gran precursor d'aquest canvi en la forma de joc s'ha demostrat que ha estat l'equip dels Houston Rockets, des de l'entrada de Daryl Morey i l'estudi de les dades massives per millorar el rendiment va determinar que llançant més de 3 punts obtindrien més victòries. S'ha vist que durant l'última dècada van ser el primer equip en començar a augmentar el nombre de llançaments intentats de 3 punts i assolir un percentatge de victòries més elevat a mesura que avançaven les temporades. Altres equips han anat copiant el model i augmentant temporada rere temporada els llançaments exteriors, fins arribar a la NBA que hi ha actualment (d'aquí que a les últimes dues temporades el creixement s'hagi estancat, com es mostra en els gràfics).

Aquest canvi de tendència en el llançament exterior ha implicat grans canvis en la forma de jugar dels equips en la NBA i de retruc en el món del bàsquet, el joc es centra cada cop més en intentar finalitzar o amb esmaixada (finalitzar deixant la pilota just en la cistella) o amb llançament exterior. Aquest fet, provoca que cada cop els jugadors siguin més atlètics i tècnicament dotats per llançar des de la línia exterior. Els jugadors interiors que fa una dècada eren més lents i amb poca capacitat de llançament exterior, han passat a ser jugadors més versàtils, amb agilitat, capacitat tècnica i amb capacitat de llançar des de l'exterior. Aquests canvis també provoquen que les defenses s'hagin vist condicionades a treballar més la part exterior.

La NBA és un referent per a la resta del món del bàsquet ja que tenen els millors jugadors, per tant, sorgeix el dubte de si a nivell de lligues europees de bàsquet o a petita escala en el nostre propi bàsquet amateur es veuria aquest canvi de tendència en el llançament exterior degut a la còpia dels patrons de la NBA.

Així doncs, es podria continuar la investigació comprovant el mateix que s'ha fet en el llançament exterior per a la lliga europea de bàsquet i pel bàsquet amateur local. En cas afirmatiu en ambdós casos, es podria determinar si hi ha una relació directa entre la part professional i l'amateur, és a dir, si els canvis que es van assolint en la part professional deriven a que en petita escala s'intentin imitar.

Però no només el llançament exterior és el que s'ha vist modificat, els equips NBA estan invertint cada cop més en analitzar la gran quantitat de dades que poden obtenir dels partits, morfologia dels jugadors, prevenció de lesions etc. Aquest anàlisi pot provocar que la NBA d'aquí 10 anys sigui completament diferent a la que tenim actualment. De fet, anualment ja es van veient modificacions constants en el propi joc dels equips que volen treure avantatge de qualsevol circumstància. Per tant, també es podria continuar la investigació intentat comprovar aquests canvis en el període analitzat. Per exemple, en l'Estat de l'Art s'ha esmentat el concepte *GHOSTING* i el *Canvi de realitat a l'hora de seleccionar jugadors en el Draft*, seria interessant poder fer una comprovació dels equips que estan aplicant aquests mètodes i en quines implicacions estan derivant.

De la mateixa manera que l'anàlisi realitzat es pot traslladar en l'estudi de múltiples aspectes del bàsquet, com s'ha comentat, també es podria traslladar a d'altres esports on l'estudi del Big Data està en creixement i pot aportar beneficis a nivell de millora de rendiment i evolució del propi esport. La lliga de Nacional de Futbol Americana (NFL), la lliga Nacional de Beisbol Americana (MLS) o el ciclisme en la UCI World Tour, són possibles lligues a analitzar ja que els seus respectius equips utilitzen l'aplicació de les dades constantment per millorar els resultats.

## 7. BIBLIOGRAFIA

- **Llibres.**

Michael Lewis (2004). Moneyball: El arte de ganar en un juego. W Norton & Co.

Paola Zuccolotto (2017). Basketball data science, Chaoman & Hall/CRC Data Science Series.

Sergio J.Ibañez Godoy (2018). Una mirada hacia la investigación e innovación sobre el baloncesto. Universidad de Extremadura.

- **Articles científics.**

Thomas Seidl (2018), Aditya Cherukumudi, Andrew Hartnett, Peter Carr, Patrick Lucey. Bhostgusters: Realtime Interactive Play Sketching with Syntehsized NBA Defenders.

Daniel Sailofsky (2018). Drafting Errors and Decision Making Bias in the NBA Draft.

Nazanin Mehrasa (2018). Deep Learning of Player Trajectory Representations for Team Activity Analysis.

Jonathan Skaza (2018). The Advantage of Doubling: A Deep Reinforcement Learning Approach to Studying the Double Team in the NBA.

Articles disponibles a:

[https://www.sloansportsconference.com/activities/research-papers/2018\\_research-paper-finalists-posters](https://www.sloansportsconference.com/activities/research-papers/2018_research-paper-finalists-posters).

- **Pàgines webs.**

Revista amb articles de bàsquet SOLOBASKET. [Consulta: 3 de març del 2023].

Disponible a <https://solobasket.com>.

Blogs d'avançaments científics en diferents areas. [Consulta: 1 d'abril del 2023].

Disponible a <https://towardsdatascience.com/>.

Pàgina web oficial de la NBA per les dades. [Consulta: 2 de febrer del 2023].

Disponible a <https://www.nba.com/stats/>.

Pàgina web per contrastar i afegir dades. [Consulta: 3 de Febrer del 2023].

Disponible a [basketball.realgm.com](https://basketball.realgm.com).

Startup amb estadístiques en directe. [Consulta: 1 de març del 2023].

Disponible a [NBN23.com/es/](https://NBN23.com/es/).

Apunts de la Universitat de Barcelona en Grau Estadística. [Consulta: 15 de Gener del 2023].

Pàgina web per contrastar i obtindrà informació extra del codi R que es vol aplicar. [Consulta: 3 de Marc del 2023]. Disponible a <https://www.r-coder.com>.

Pàgina web per contrastar i obtenir informació extra del codi SAS que es vol aplicar. SAS Help Center. [Consulta: 3 de Març del 2023].

Disponible a <https://www.sas.com>.



- **Vídeos.**

42 Analytics. (2014) SSAC14: SSAC14: BASKETBALL Analytics. [Consulta: 4 d'abril del 2023]

Disponible a : <https://www.youtube.com/watch?v=aZaCcg20TuY>.

42 Analytics. (2018) SSAC18: Take That for Data: Basketball Analytics. [Consulta: 4 d'abril del 2023].

Disponible a : <https://www.youtube.com/watch?v=jL2ZSMQOOS4&t=945s>.

## 8. ANNEX

- **Codi SAS**

*/\*Importar dades de zona de llançament\*/*

```
proc import out= temp1 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular.xlsx"
  dbms=excel replace ;
  sheet="01_02";
  getnames=yes;
run;
proc import out= temp2 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular.xlsx"
  dbms=excel replace ;
  sheet="02_03";
  getnames=yes;
run;
proc import out= temp3 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular.xlsx"
  dbms=excel replace ;
  sheet="03_04";
  getnames=yes;
run;
proc import out= temp4 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular.xlsx"
  dbms=excel replace ;
  sheet="04_05";
  getnames=yes;
run;
proc import out= temp5 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular.xlsx"
  dbms=excel replace ;
  sheet="05_06";
  getnames=yes;
run;
proc import out= temp6 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular.xlsx"
  dbms=excel replace ;
  sheet="06_07";
  getnames=yes;
run;
proc import out= temp7 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular.xlsx"
  dbms=excel replace ;
  sheet="07_08";
  getnames=yes;
run;
proc import out= temp8 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular.xlsx"
  dbms=excel replace ;
  sheet="08_09";
  getnames=yes;
run;
proc import out= temp9 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular.xlsx"
  dbms=excel replace ;
  sheet="09_10";
  getnames=yes;
run;
proc import out= temp10 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular.xlsx"
  dbms=excel replace ;
  sheet="10_11";
  getnames=yes;
run;
proc import out= temp11 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular.xlsx"
  dbms=excel replace ;
  sheet="11_12";
  getnames=yes;
run;
proc import out= temp12 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular.xlsx"
  dbms=excel replace ;
  sheet="12_13";
  getnames=yes;
run;
```

```

proc import out= temp13 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular.xlsx"
  dbms=excel replace ;
  sheet="13_14";
  getnames=yes;
run;
proc import out= temp14 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular.xlsx"
  dbms=excel replace ;
  sheet="14_15";
  getnames=yes;
run;
proc import out= temp15 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular.xlsx"
  dbms=excel replace ;
  sheet="15_16";
  getnames=yes;
run;
proc import out= temp16 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular.xlsx"
  dbms=excel replace ;
  sheet="16_17";
  getnames=yes;
run;
proc import out= temp17 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular.xlsx"
  dbms=excel replace ;
  sheet="17_18";
  getnames=yes;
run;
proc import out= temp18 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular.xlsx"
  dbms=excel replace ;
  sheet="18_19";
  getnames=yes;
run;
proc import out= temp19 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular.xlsx"
  dbms=excel replace ;
  sheet="19_20";
  getnames=yes;
run;
proc import out= temp20 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular.xlsx"
  dbms=excel replace ;
  sheet="20_21";
  getnames=yes;
run;
proc import out= temp21 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular.xlsx"
  dbms=excel replace ;
  sheet="21_22";
  getnames=yes;
run;

/*Playoffs*/
proc import out= playoff1 datafile= "c:\proves\secret_project\dades\tirs_evolucion_playof.xlsx"
  dbms=excel replace ;
  sheet="01_02";
  getnames=yes;
run;
proc import out= playoff2 datafile= "c:\proves\secret_project\dades\tirs_evolucion_playof.xlsx"
  dbms=excel replace ;
  sheet="02_03";
  getnames=yes;
run;
proc import out= playoff3 datafile= "c:\proves\secret_project\dades\tirs_evolucion_playof.xlsx"
  dbms=excel replace ;
  sheet="03_04";
  getnames=yes;
run;
proc import out= playoff4 datafile= "c:\proves\secret_project\dades\tirs_evolucion_playof.xlsx"
  dbms=excel replace ;
  sheet="04_05";
  getnames=yes;
run;

```

```

proc import out= playof5 datafile= "c:\proves\secret_project\dades\tirs_evolucion_playof.xlsx"
    dbms=excel replace ;
    sheet="05_06";
    getnames=yes;
run;
proc import out= playof6 datafile= "c:\proves\secret_project\dades\tirs_evolucion_playof.xlsx"
    dbms=excel replace ;
    sheet="06_07";
    getnames=yes;
run;
proc import out= playof7 datafile= "c:\proves\secret_project\dades\tirs_evolucion_playof.xlsx"
    dbms=excel replace ;
    sheet="07_08";
    getnames=yes;
run;
proc import out= playof8 datafile= "c:\proves\secret_project\dades\tirs_evolucion_playof.xlsx"
    dbms=excel replace ;
    sheet="08_09";
    getnames=yes;
run;
proc import out= playof9 datafile= "c:\proves\secret_project\dades\tirs_evolucion_playof.xlsx"
    dbms=excel replace ;
    sheet="09_10";
    getnames=yes;
run;
proc import out= playof10 datafile= "c:\proves\secret_project\dades\tirs_evolucion_playof.xlsx"
    dbms=excel replace ;
    sheet="10_11";
    getnames=yes;
run;
proc import out= playof11 datafile= "c:\proves\secret_project\dades\tirs_evolucion_playof.xlsx"
    dbms=excel replace ;
    sheet="11_12";
    getnames=yes;
run;
proc import out= playof12 datafile= "c:\proves\secret_project\dades\tirs_evolucion_playof.xlsx"
    dbms=excel replace ;
    sheet="12_13";
    getnames=yes;
run;
proc import out= playof13 datafile= "c:\proves\secret_project\dades\tirs_evolucion_playof.xlsx"
    dbms=excel replace ;
    sheet="13_14";
    getnames=yes;
run;
proc import out= playof14 datafile= "c:\proves\secret_project\dades\tirs_evolucion_playof.xlsx"
    dbms=excel replace ;
    sheet="14_15";
    getnames=yes;
run;
proc import out= playof15 datafile= "c:\proves\secret_project\dades\tirs_evolucion_playof.xlsx"
    dbms=excel replace ;
    sheet="15_16";
    getnames=yes;
run;
proc import out= playof16 datafile= "c:\proves\secret_project\dades\tirs_evolucion_playof.xlsx"
    dbms=excel replace ;
    sheet="16_17";
    getnames=yes;
run;
proc import out= playof17 datafile= "c:\proves\secret_project\dades\tirs_evolucion_playof.xlsx"
    dbms=excel replace ;
    sheet="17_18";
    getnames=yes;
run;
proc import out= playof18 datafile= "c:\proves\secret_project\dades\tirs_evolucion_playof.xlsx"
    dbms=excel replace ;

```

```

        sheet="18_19";
        getnames=yes;
run;
proc import out= playof19 datafile= "c:\proves\secret_project\dades\tirs_evolutio_playof.xlsx"
    dbms=excel replace ;
    sheet="19_20";
    getnames=yes;
run;
proc import out= playof20 datafile= "c:\proves\secret_project\dades\tirs_evolutio_playof.xlsx"
    dbms=excel replace ;
    sheet="20_21";
    getnames=yes;
run;
proc import out= playof21 datafile= "c:\proves\secret_project\dades\tirs_evolutio_playof.xlsx"
    dbms=excel replace ;
    sheet="21_22";
    getnames=yes;
run;

/*Importar dades de tirs lliures*/

proc import out= stats1 datafile= "c:\proves\secret_project\dades\tirs_lliuers.xlsx"
    dbms=excel replace ;
    sheet="2001";
    getnames=yes;
run;
proc import out= stats2 datafile= "c:\proves\secret_project\dades\tirs_lliuers.xlsx"
    dbms=excel replace ;
    sheet="2002";
    getnames=yes;
run;
proc import out= stats3 datafile= "c:\proves\secret_project\dades\tirs_lliuers.xlsx"
    dbms=excel replace ;
    sheet="2003";
    getnames=yes;
run;
proc import out= stats4 datafile= "c:\proves\secret_project\dades\tirs_lliuers.xlsx"
    dbms=excel replace ;
    sheet="2004";
    getnames=yes;
run;
proc import out= stats5 datafile= "c:\proves\secret_project\dades\tirs_lliuers.xlsx"
    dbms=excel replace ;
    sheet="2005";
    getnames=yes;
run;
proc import out= stats6 datafile= "c:\proves\secret_project\dades\tirs_lliuers.xlsx"
    dbms=excel replace ;
    sheet="2006";
    getnames=yes;
run;
proc import out= stats7 datafile= "c:\proves\secret_project\dades\tirs_lliuers.xlsx"
    dbms=excel replace ;
    sheet="2007";
    getnames=yes;
run;
proc import out= stats8 datafile= "c:\proves\secret_project\dades\tirs_lliuers.xlsx"
    dbms=excel replace ;
    sheet="2008";
    getnames=yes;
run;
proc import out= stats9 datafile= "c:\proves\secret_project\dades\tirs_lliuers.xlsx"
    dbms=excel replace ;
    sheet="2009";
    getnames=yes;
run;

```

```

proc import out= stats10 datafile= "c:\proves\secret_project\dades\tirs_lliures.xlsx"
  dbms=excel replace ;
  sheet="2010";
  getnames=yes;
run;
proc import out= stats11 datafile= "c:\proves\secret_project\dades\tirs_lliures.xlsx"
  dbms=excel replace ;
  sheet="2011";
  getnames=yes;
run;
proc import out= stats12 datafile= "c:\proves\secret_project\dades\tirs_lliures.xlsx"
  dbms=excel replace ;
  sheet="2012";
  getnames=yes;
run;
proc import out= stats13 datafile= "c:\proves\secret_project\dades\tirs_lliures.xlsx"
  dbms=excel replace ;
  sheet="2013";
  getnames=yes;
run;
proc import out= stats14 datafile= "c:\proves\secret_project\dades\tirs_lliures.xlsx"
  dbms=excel replace ;
  sheet="2014";
  getnames=yes;
run;
proc import out= stats15 datafile= "c:\proves\secret_project\dades\tirs_lliures.xlsx"
  dbms=excel replace ;
  sheet="2015";
  getnames=yes;
run;
proc import out= stats16 datafile= "c:\proves\secret_project\dades\tirs_lliures.xlsx"
  dbms=excel replace ;
  sheet="2016";
  getnames=yes;
run;
proc import out= stats17 datafile= "c:\proves\secret_project\dades\tirs_lliures.xlsx"
  dbms=excel replace ;
  sheet="2017";
  getnames=yes;
run;
proc import out= stats18 datafile= "c:\proves\secret_project\dades\tirs_lliures.xlsx"
  dbms=excel replace ;
  sheet="2018";
  getnames=yes;
run;
proc import out= stats19 datafile= "c:\proves\secret_project\dades\tirs_lliures.xlsx"
  dbms=excel replace ;
  sheet="2019";
  getnames=yes;
run;
proc import out= stats20 datafile= "c:\proves\secret_project\dades\tirs_lliures.xlsx"
  dbms=excel replace ;
  sheet="2020";
  getnames=yes;
run;
proc import out= stats21 datafile= "c:\proves\secret_project\dades\tirs_lliures.xlsx"
  dbms=excel replace ;
  sheet="2021";
  getnames=yes;
run;

/*Importar dades de zona de llançament en la posició d'aler*/

proc import out= forw1 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular_forward.xlsx"
  dbms=excel replace ;
  sheet="01_02";

```

```

        getnames=yes;
run;
proc import out= forw2 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular_forward.xlsx"
    dbms=excel replace ;
    sheet="02_03";
    getnames=yes;
run;
proc import out= forw3 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular_forward.xlsx"
    dbms=excel replace ;
    sheet="03_04";
    getnames=yes;
run;
proc import out= forw4 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular_forward.xlsx"
    dbms=excel replace ;
    sheet="04_05";
    getnames=yes;
run;
proc import out= forw5 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular_forward.xlsx"
    dbms=excel replace ;
    sheet="05_06";
    getnames=yes;
run;
proc import out= forw6 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular_forward.xlsx"
    dbms=excel replace ;
    sheet="06_07";
    getnames=yes;
run;
proc import out= forw7 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular_forward.xlsx"
    dbms=excel replace ;
    sheet="07_08";
    getnames=yes;
run;
proc import out= forw8 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular_forward.xlsx"
    dbms=excel replace ;
    sheet="08_09";
    getnames=yes;
run;
proc import out= forw9 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular_forward.xlsx"
    dbms=excel replace ;
    sheet="09_10";
    getnames=yes;
run;
proc import out= forw10 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular_forward.xlsx"
    dbms=excel replace ;
    sheet="10_11";
    getnames=yes;
run;
proc import out= forw11 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular_forward.xlsx"
    dbms=excel replace ;
    sheet="11_12";
    getnames=yes;
run;
proc import out= forw12 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular_forward.xlsx"
    dbms=excel replace ;
    sheet="12_13";
    getnames=yes;
run;
proc import out= forw13 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular_forward.xlsx"
    dbms=excel replace ;
    sheet="13_14";
    getnames=yes;
run;
proc import out= forw14 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular_forward.xlsx"
    dbms=excel replace ;
    sheet="14_15";
    getnames=yes;
run;

```

```

proc import out= forw15 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular_forward.xlsx"
  dbms=excel replace ;
  sheet="15_16";
  getnames=yes;
run;
proc import out= forw16 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular_forward.xlsx"
  dbms=excel replace ;
  sheet="16_17";
  getnames=yes;
run;
proc import out= forw17 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular_forward.xlsx"
  dbms=excel replace ;
  sheet="17_18";
  getnames=yes;
run;
proc import out= forw18 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular_forward.xlsx"
  dbms=excel replace ;
  sheet="18_19";
  getnames=yes;
run;
proc import out= forw19 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular_forward.xlsx"
  dbms=excel replace ;
  sheet="19_20";
  getnames=yes;
run;
proc import out= forw20 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular_forward.xlsx"
  dbms=excel replace ;
  sheet="20_21";
  getnames=yes;
run;
proc import out= forw21 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular_forward.xlsx"
  dbms=excel replace ;
  sheet="21_22";
  getnames=yes;
run;

/*Importar dades de zona de llançament en la posició de pivot*/

proc import out= center1 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular_center.xlsx"
  dbms=excel replace ;
  sheet="01_02";
  getnames=yes;
run;
proc import out= center2 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular_center.xlsx"
  dbms=excel replace ;
  sheet="02_03";
  getnames=yes;
run;
proc import out= center3 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular_center.xlsx"
  dbms=excel replace ;
  sheet="03_04";
  getnames=yes;
run;
proc import out= center4 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular_center.xlsx"
  dbms=excel replace ;
  sheet="04_05";
  getnames=yes;
run;
proc import out= center5 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular_center.xlsx"
  dbms=excel replace ;
  sheet="05_06";
  getnames=yes;
run;
proc import out= center6 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular_center.xlsx"
  dbms=excel replace ;
  sheet="06_07";
  getnames=yes;

```



```

run;
proc import out= center7 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular_center.xlsx"
  dbms=excel replace ;
  sheet="07_08";
  getnames=yes;

run;
proc import out= center8 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular_center.xlsx"
  dbms=excel replace ;
  sheet="08_09";
  getnames=yes;

run;
proc import out= center9 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular_center.xlsx"
  dbms=excel replace ;
  sheet="09_10";
  getnames=yes;

run;
proc import out= center10 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular_center.xlsx"
  dbms=excel replace ;
  sheet="10_11";
  getnames=yes;

run;
proc import out= center11 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular_center.xlsx"
  dbms=excel replace ;
  sheet="11_12";
  getnames=yes;

run;
proc import out= center12 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular_center.xlsx"
  dbms=excel replace ;
  sheet="12_13";
  getnames=yes;

run;
proc import out= center13 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular_center.xlsx"
  dbms=excel replace ;
  sheet="13_14";
  getnames=yes;

run;
proc import out= center14 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular_center.xlsx"
  dbms=excel replace ;
  sheet="14_15";
  getnames=yes;

run;
proc import out= center15 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular_center.xlsx"
  dbms=excel replace ;
  sheet="15_16";
  getnames=yes;

run;
proc import out= center16 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular_center.xlsx"
  dbms=excel replace ;
  sheet="16_17";
  getnames=yes;

run;
proc import out= center17 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular_center.xlsx"
  dbms=excel replace ;
  sheet="17_18";
  getnames=yes;

run;
proc import out= center18 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular_center.xlsx"
  dbms=excel replace ;
  sheet="18_19";
  getnames=yes;

run;
proc import out= center19 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular_center.xlsx"
  dbms=excel replace ;
  sheet="19_20";
  getnames=yes;

run;
proc import out= center20 datafile= "c:\proves\secret_project\dades\tirs_evolucion_regular_center.xlsx"

```

```

        dbms=excel replace ;
        sheet="20_21";
        getnames=yes;
run;
proc import out= center21 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular_center.xlsx"
        dbms=excel replace ;
        sheet="21_22";
        getnames=yes;
run;

/*Importar dades de zona de llançament en la posició de base*/

proc import out= guard1 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular_guard.xlsx"
        dbms=excel replace ;
        sheet="01_02";
        getnames=yes;
run;
proc import out= guard2 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular_guard.xlsx"
        dbms=excel replace ;
        sheet="02_03";
        getnames=yes;
run;
proc import out= guard3 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular_guard.xlsx"
        dbms=excel replace ;
        sheet="03_04";
        getnames=yes;
run;
proc import out= guard4 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular_guard.xlsx"
        dbms=excel replace ;
        sheet="04_05";
        getnames=yes;
run;
proc import out= guard5 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular_guard.xlsx"
        dbms=excel replace ;
        sheet="05_06";
        getnames=yes;
run;
proc import out= guard6 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular_guard.xlsx"
        dbms=excel replace ;
        sheet="06_07";
        getnames=yes;
run;
proc import out= guard7 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular_guard.xlsx"
        dbms=excel replace ;
        sheet="07_08";
        getnames=yes;
run;
proc import out= guard8 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular_guard.xlsx"
        dbms=excel replace ;
        sheet="08_09";
        getnames=yes;
run;
proc import out= guard9 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular_guard.xlsx"
        dbms=excel replace ;
        sheet="09_10";
        getnames=yes;
run;
proc import out= guard10 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular_guard.xlsx"
        dbms=excel replace ;
        sheet="10_11";
        getnames=yes;
run;
proc import out= guard11 datafile= "c:\proves\secret_project\dades\tirs_evolutio_regular_guard.xlsx"
        dbms=excel replace ;
        sheet="11_12";
        getnames=yes;
run;

```

```

proc import out= guard12 datafile= "c:\proves\secret_project\dades\tirs_evolicio_regular_guard.xlsx"
    dbms=excel replace ;
    sheet="12_13";
    getnames=yes;
run;
proc import out= guard13 datafile= "c:\proves\secret_project\dades\tirs_evolicio_regular_guard.xlsx"
    dbms=excel replace ;
    sheet="13_14";
    getnames=yes;
run;
proc import out= guard14 datafile= "c:\proves\secret_project\dades\tirs_evolicio_regular_guard.xlsx"
    dbms=excel replace ;
    sheet="14_15";
    getnames=yes;
run;
proc import out= guard15 datafile= "c:\proves\secret_project\dades\tirs_evolicio_regular_guard.xlsx"
    dbms=excel replace ;
    sheet="15_16";
    getnames=yes;
run;
proc import out= guard16 datafile= "c:\proves\secret_project\dades\tirs_evolicio_regular_guard.xlsx"
    dbms=excel replace ;
    sheet="16_17";
    getnames=yes;
run;
proc import out= guard17 datafile= "c:\proves\secret_project\dades\tirs_evolicio_regular_guard.xlsx"
    dbms=excel replace ;
    sheet="17_18";
    getnames=yes;
run;
proc import out= guard18 datafile= "c:\proves\secret_project\dades\tirs_evolicio_regular_guard.xlsx"
    dbms=excel replace ;
    sheet="18_19";
    getnames=yes;
run;
proc import out= guard19 datafile= "c:\proves\secret_project\dades\tirs_evolicio_regular_guard.xlsx"
    dbms=excel replace ;
    sheet="19_20";
    getnames=yes;
run;
proc import out= guard20 datafile= "c:\proves\secret_project\dades\tirs_evolicio_regular_guard.xlsx"
    dbms=excel replace ;
    sheet="20_21";
    getnames=yes;
run;
proc import out= guard21 datafile= "c:\proves\secret_project\dades\tirs_evolicio_regular_guard.xlsx"
    dbms=excel replace ;
    sheet="21_22";
    getnames=yes;
run;

/*Importar dades de percentatge de victoria*/

proc import out= victoria1 datafile= "c:\proves\secret_project\dades\victories.xlsx"
    dbms=excel replace ;
    sheet="01_02";
    getnames=yes;
run;
proc import out= victoria2 datafile= "c:\proves\secret_project\dades\victories.xlsx"
    dbms=excel replace ;
    sheet="02_03";
    getnames=yes;
run;
proc import out= victoria3 datafile= "c:\proves\secret_project\dades\victories.xlsx"
    dbms=excel replace ;
    sheet="03_04";
    getnames=yes;

```

```

run;
proc import out= victoria4 datafile= "c:\proves\secret_project\dades\victories.xlsx"
  dbms=excel replace ;
  sheet="04_05";
  getnames=yes;

run;
proc import out= victoria5 datafile= "c:\proves\secret_project\dades\victories.xlsx"
  dbms=excel replace ;
  sheet="05_06";
  getnames=yes;

run;
proc import out= victoria6 datafile= "c:\proves\secret_project\dades\victories.xlsx"
  dbms=excel replace ;
  sheet="06_07";
  getnames=yes;

run;
proc import out= victoria7 datafile= "c:\proves\secret_project\dades\victories.xlsx"
  dbms=excel replace ;
  sheet="07_08";
  getnames=yes;

run;
proc import out= victoria8 datafile= "c:\proves\secret_project\dades\victories.xlsx"
  dbms=excel replace ;
  sheet="08_09";
  getnames=yes;

run;
proc import out= victoria9 datafile= "c:\proves\secret_project\dades\victories.xlsx"
  dbms=excel replace ;
  sheet="09_10";
  getnames=yes;

run;
proc import out= victoria10 datafile= "c:\proves\secret_project\dades\victories.xlsx"
  dbms=excel replace ;
  sheet="10_11";
  getnames=yes;

run;
proc import out= victoria11 datafile= "c:\proves\secret_project\dades\victories.xlsx"
  dbms=excel replace ;
  sheet="11_12";
  getnames=yes;

run;
proc import out= victoria12 datafile= "c:\proves\secret_project\dades\victories.xlsx"
  dbms=excel replace ;
  sheet="12_13";
  getnames=yes;

run;
proc import out= victoria13 datafile= "c:\proves\secret_project\dades\victories.xlsx"
  dbms=excel replace ;
  sheet="13_14";
  getnames=yes;

run;
proc import out= victoria14 datafile= "c:\proves\secret_project\dades\victories.xlsx"
  dbms=excel replace ;
  sheet="14_15";
  getnames=yes;

run;
proc import out= victoria15 datafile= "c:\proves\secret_project\dades\victories.xlsx"
  dbms=excel replace ;
  sheet="15_16";
  getnames=yes;

run;
proc import out= victoria16 datafile= "c:\proves\secret_project\dades\victories.xlsx"
  dbms=excel replace ;
  sheet="16_17";
  getnames=yes;

run;
proc import out= victoria17 datafile= "c:\proves\secret_project\dades\victories.xlsx"

```

```

        dbms=excel replace ;
        sheet="17_18";
        getnames=yes;
run;
proc import out= victoria18 datafile= "c:\proves\secret_project\dades\victories.xlsx"
        dbms=excel replace ;
        sheet="18_19";
        getnames=yes;
run;
proc import out= victoria19 datafile= "c:\proves\secret_project\dades\victories.xlsx"
        dbms=excel replace ;
        sheet="19_20";
        getnames=yes;
run;
proc import out= victoria20 datafile= "c:\proves\secret_project\dades\victories.xlsx"
        dbms=excel replace ;
        sheet="20_21";
        getnames=yes;
run;
proc import out= victoria21 datafile= "c:\proves\secret_project\dades\victories.xlsx"
        dbms=excel replace ;
        sheet="21_22";
        getnames=yes;
run;

/*Adjuntar les diferents temporades en un sol data*/
data regul;
set temp;;
run;
proc sort data=regul; by temporada; run;

data play;
set playof;;
run;
proc sort data=play; by temporada; run;

data stats;
set stats;;
run;

data forward;
set forw;;
run;

data center;
set center;;
run;

data guard;
set guard;;
run;

data victoria;
set victoria;;
run;

/*Afegir càlculs de variables que no es tenen*/
data regular;
set regul;
        T3_FGM = CO_FGM + ATB_FGM;
        T3_FGA = CO_FGA + ATB_FGA;
        T3_FG = round((T3_FGM/T3_FGA)*100,.1);
        TTE=RA_FGM+IP_FGM+MR_FGM+CO_FGM+ATB_FGM;
        TTI=RA_FGA+IP_FGA+MR_FGA+CO_FGA+ATB_FGA;
        TTP=round((TTE/TTI)*100,.1);
run;

```

```

data forward;
set forward;
    T3_FGM = CO_FGM + ATB_FGM;
    T3_FGA = CO_FGA + ATB_FGA;
    T3_FG = round((T3_FGM/T3_FGA)*100,.1);
    TTE=RA_FGM+IP_FGM+MR_FGM+CO_FGM+ATB_FGM;
    TTI=RA_FGA+IP_FGA+MR_FGA+CO_FGA+ATB_FGA;
    TTP=round((TTE/TTI)*100,.1);

run;

data center;
set center;
    T3_FGM = CO_FGM + ATB_FGM;
    T3_FGA = CO_FGA + ATB_FGA;
    T3_FG = round((T3_FGM/T3_FGA)*100,.1);
    TTE=RA_FGM+IP_FGM+MR_FGM+CO_FGM+ATB_FGM;
    TTI=RA_FGA+IP_FGA+MR_FGA+CO_FGA+ATB_FGA;
    TTP=round((TTE/TTI)*100,.1);

run;

data guard;
set guard;
    T3_FGM = CO_FGM + ATB_FGM;
    T3_FGA = CO_FGA + ATB_FGA;
    T3_FG = round((T3_FGM/T3_FGA)*100,.1);
    TTE=RA_FGM+IP_FGM+MR_FGM+CO_FGM+ATB_FGM;
    TTI=RA_FGA+IP_FGA+MR_FGA+CO_FGA+ATB_FGA;
    TTP=round((TTE/TTI)*100,.1);

run;

data finales;
set play;
    T3_FGM = CO_FGM + ATB_FGM;
    T3_FGA = CO_FGA + ATB_FGA;
    T3_FG = round((T3_FGM/T3_FGA)*100,.1);

run;

/*Gràfic boxplot*/

PROC SGPLOT DATA = regular;
    VBOX T3_FGA
    / category = TEAM;
title 'Llançaments a cistella de mitjana en el període estudiat per equips';
run;

/*Renombrem els equips que s'han vist modificats en el pas de les temporades*/

data stats;
set stats;
if team = "Charlotte Bobcats" then team= "Charlotte Hornets";
if team = "LA Clippers" then team= "Los Angeles Clippers";
if team = "New Jersey Nets" then team="Brooklyn Nets";
if team = "New Orleans Hornets" then team="New Orleans Pelicans";
if team = "New Orleans/Oklahoma C" then team="New Orleans Pelicans";
if team = "Seattle SuperSonics" then team="Oklahoma City Thunder" ;
run;

data center;
set center;
if team = "Charlotte Bobcats" then team= "Charlotte Hornets";
if team = "LA Clippers" then team= "Los Angeles Clippers";
if team = "New Jersey Nets" then team="Brooklyn Nets";
if team = "New Orleans Hornets" then team="New Orleans Pelicans";
if team = "New Orleans/Oklahoma C" then team="New Orleans Pelicans";
if team = "Seattle SuperSonics" then team="Oklahoma City Thunder" ;
run;

```

```

data guard;
set guard;
if team = "Charlotte Bobcats" then team= "Charlotte Hornets";
if team = "LA Clippers" then team= "Los Angeles Clippers";
if team = "New Jersey Nets" then team="Brooklyn Nets";
if team = "New Orleans Hornets" then team="New Orleans Pelicans";
if team = "New Orleans/Oklahoma C" then team="New Orleans Pelicans";
if team = "Seattle SuperSonics" then team="Oklahoma City Thunder" ;
run;

```

```

data forward;
set forward;
if team = "Charlotte Bobcats" then team= "Charlotte Hornets";
if team = "LA Clippers" then team= "Los Angeles Clippers";
if team = "New Jersey Nets" then team="Brooklyn Nets";
if team = "New Orleans Hornets" then team="New Orleans Pelicans";
if team = "New Orleans/Oklahoma C" then team="New Orleans Pelicans";
if team = "Seattle SuperSonics" then team="Oklahoma City Thunder" ;
run;

```

```

data final;
set final;
if team = "Charlotte Bobcats" then team= "Charlotte Hornets";
if team = "LA Clippers" then team= "Los Angeles Clippers";
if team = "New Jersey Nets" then team="Brooklyn Nets";
if team = "New Orleans Hornets" then team="New Orleans Pelicans";
if team = "New Orleans/Oklahoma C" then team="New Orleans Pelicans";
if team = "Seattle SuperSonics" then team="Oklahoma City Thunder" ;
run;

```

/\*Es defineixen els diferents clusters que s'han obtingut en el codi R\*/

```

data cluster;
set regular;
if team = "Atlanta Hawks" then cluster="G1" ;
if team = "Boston Celtics" then cluster="G1";
if team = "Brooklyn Nets" then cluster="G1";
if team = "Charlotte Hornets" then cluster="G1";
if team = "Cleveland Cavaliers" then cluster="G1";
if team = "Denver Nuggets" then cluster="G1";
if team = "Indiana Pacers" then cluster="G1";
if team = "Los Angeles Lakers" then cluster="G1";
if team = "Milwaukee Bucks" then cluster="G1";
if team = "New York Knicks" then cluster="G1";
if team = "Oklahoma City Thunder" then cluster="G1";
if team = "Orlando Magic" then cluster="G1";
if team = "Phoenix Suns" then cluster="G1";
if team = "Portland Trail Blazers" then cluster="G1";
if team = "Chicago Bulls" then cluster="G2";
if team = "Detroit Pistons" then cluster="G2";
if team = "Los Angeles Clippers" then cluster="G2";
if team = "Memphis Grizzlies" then cluster="G2";
if team = "Minnesota Timberwolves" then cluster="G2";
if team = "New Orleans Pelicans" then cluster="G2";
if team = "Philadelphia 76ers" then cluster="G2";
if team = "Utah Jazz" then cluster="G2";
if team = "Washington Wizards" then cluster="G2";
if team = "Dallas Mavericks" then cluster="G3";
if team = "Golden State Warriors" then cluster="G3";
if team = "Houston Rockets" then cluster="G3";
if team = "Miami Heat" then cluster="G1";
if team = "Sacramento Kings" then cluster="G2";
if team = "San Antonio Spurs" then cluster="G2";
if team = "Toronto Raptors" then cluster="G1";
run;

```

/\*Calculem les bases de dades mitjançant tabulacions utilitzant SQL \*/

```
proc sql;
  create table comp_tirs_lliuers as
  select temporada, mean(FGM) as FGM, mean(FGA) as FGA, mean(FG_) as FG_, mean(T3_PM) as T3_PM,
  mean(T3_PA) as T3_PA, mean(T3_P_) as T3_P_, mean(FTM) as FTM,
  mean(FTA) as FTA, mean(FT_) as FT_
  from stats
  group by temporada;
quit;
```

```
proc sql;
  create table comp_guard as
  select temporada, mean(T3_FGM) as T3_FGM, mean(T3_FGA) as T3_FGA, mean(T3_FG) as T3_FG,
  mean(CO_FGM) as CO_FGM, mean(CO_FGA) as CO_FGA, mean(CO_FG_) as CO_FG_,
  mean(ATB_FGM) as ATB_FGM, mean(ATB_FGA) as ATB_FGA, mean(ATB_FG_) as ATB_FG_,
  mean(RA_FGM) as RA_FGM, mean(RA_FGA) as RA_FGA,
  mean(RA_FG_) as RA_FG_, mean(IP_FGM) as IP_FGM, mean(IP_FGA) as IP_FGA,
  mean(IP_FG_) as IP_FG_, mean(MR_FGM) as MR_FGM, mean(MR_FGA) as MR_FGA,
  mean(MR_FG_) as MR_FG_
  from guard
  group by temporada;
quit;
```

```
proc sql;
  create table comp_forward as
  select temporada, mean(T3_FGM) as T3_FGM, mean(T3_FGA) as T3_FGA, mean(T3_FG) as T3_FG,
  mean(CO_FGM) as CO_FGM, mean(CO_FGA) as CO_FGA, mean(CO_FG_) as CO_FG_,
  mean(ATB_FGM) as ATB_FGM, mean(ATB_FGA) as ATB_FGA, mean(ATB_FG_) as ATB_FG_,
  mean(RA_FGM) as RA_FGM, mean(RA_FGA) as RA_FGA,
  mean(RA_FG_) as RA_FG_, mean(IP_FGM) as IP_FGM, mean(IP_FGA) as IP_FGA,
  mean(IP_FG_) as IP_FG_, mean(MR_FGM) as MR_FGM, mean(MR_FGA) as MR_FGA,
  mean(MR_FG_) as MR_FG_
  from forward
  group by temporada;
quit;
```

```
proc sql;
  create table comp_center as
  select temporada, mean(T3_FGM) as T3_FGM, mean(T3_FGA) as T3_FGA, mean(T3_FG) as T3_FG,
  mean(CO_FGM) as CO_FGM, mean(CO_FGA) as CO_FGA, mean(CO_FG_) as CO_FG_,
  mean(ATB_FGM) as ATB_FGM, mean(ATB_FGA) as ATB_FGA, mean(ATB_FG_) as ATB_FG_,
  mean(RA_FGM) as RA_FGM, mean(RA_FGA) as RA_FGA,
  mean(RA_FG_) as RA_FG_, mean(IP_FGM) as IP_FGM, mean(IP_FGA) as IP_FGA,
  mean(IP_FG_) as IP_FG_, mean(MR_FGM) as MR_FGM, mean(MR_FGA) as MR_FGA,
  mean(MR_FG_) as MR_FG_
  from center
  group by temporada;
quit;
```

```
data comp_guard;
  set comp_guard;
  posicio="guard";
```

```
run;
data comp_forward;
  set comp_forward;
  posicio="forward";
```

```
run;
data comp_center;
  set comp_center;
  posicio="center";
```

```
run;

data posicio;
```



```
set comp_guard comp_forward comp_center;
run;
```

```
proc sql;
  create table comp_total_any as
  select temporada,mean(T3_FGM) as T3_FGM, mean(T3_FGA) as T3_FGA, mean(T3_FG) as T3_FG,
  mean(CO_FGM) as CO_FGM, mean(CO_FGA) as CO_FGA, mean(CO_FG_) as CO_FG_,
  mean(ATB_FGM) as ATB_FGM, mean(ATB_FGA) as ATB_FGA, mean(ATB_FG_) as ATB_FG_,
  mean(RA_FGM) as RA_FGM, mean(RA_FGA) as RA_FGA,
  mean(RA_FG_) as RA_FG_, mean(IP_FGM) as IP_FGM, mean(IP_FGA) as IP_FGA,
  mean(IP_FG_) as IP_FG_, mean(MR_FGM) as MR_FGM, mean(MR_FGA) as MR_FGA,
  mean(MR_FG_) as MR_FG_
  from regular
  group by temporada;
quit;
```

```
proc sql;
  create table comp_total_equip as
  select team,mean(CO_FGM) as CO_FGM, mean(CO_FGA) as CO_FGA, mean(CO_FG_) as CO_FG_,
  mean(ATB_FGM) as ATB_FGM, mean(ATB_FGA) as ATB_FGA, mean(ATB_FG_) as ATB_FG_,
  mean(T3_FGM) as T3_FGM, mean(T3_FGA) as T3_FGA, mean(T3_FG) as T3_FG
  from regular
  group by team;
quit;
```

```
data test_mitjanes;
  set regular;
  if temporada in ("2001_2002","2021_2022");
run;
```

```
proc sql;
  create table definitiu as
  select *
  from
  final as a
  left join
  regular as b
  on
  a.Temporada=b.Temporada and a.Team=b.Team
  ;
```

```
quit;
```

```
proc sql;
  create table comp_total_cluster as
  select cluster,temporada,mean(T3_FGM) as T3_FGM, mean(T3_FGA) as T3_FGA, mean(T3_FG) as
  T3_FG, mean(CO_FGM) as CO_FGM, mean(CO_FGA) as CO_FGA, mean(CO_FG_) as CO_FG_,
  mean(ATB_FGM) as ATB_FGM, mean(ATB_FGA) as ATB_FGA, mean(ATB_FG_) as ATB_FG_,
  mean(RA_FGM) as RA_FGM, mean(RA_FGA) as RA_FGA,
  mean(RA_FG_) as RA_FG_, mean(IP_FGM) as IP_FGM, mean(IP_FGA) as IP_FGA,
  mean(IP_FG_) as IP_FG_, mean(MR_FGM) as MR_FGM, mean(MR_FGA) as MR_FGA,
  mean(MR_FG_) as MR_FG_
  from cluster
  group by cluster,temporada;
quit;
```

```
proc sql;
  create table definitiu_cluster as
  select cluster,temporada,mean(W) as W, mean(WIN_) as WIN_, mean(PTS) as PTS, mean(FGA) as FGA,
  mean(FG_) as FG_,
  mean(FTA) as FTA, mean(FT_) as FT_, mean(_PA) as _PA,
  mean(REB) as REB, mean(OREB) as OREB,
  mean(DREB) as DREB, mean(AST) as AST, mean(MR_FGA) as MR_FGA
  from definitiu
  group by cluster,temporada;
```

```

quit;

/*Exportació dels fitxers*/
proc export
  data=finales
  outfile="c:\proves\secret_project\dades\playof.xlsx"
  dbms=xlsx replace;
run;

proc export
  data=regular
  outfile="c:\proves\secret_project\dades\regular.xlsx"
  dbms=xlsx replace;
run;

proc export
  data=stats
  outfile="c:\proves\secret_project\dades\stats.xlsx"
  dbms=xlsx replace;
run;

proc export
  data=forward
  outfile="c:\proves\secret_project\dades\forward.xlsx"
  dbms=xlsx replace;
run;

proc export
  data=center
  outfile="c:\proves\secret_project\dades\center.xlsx"
  dbms=xlsx replace;
run;

proc export
  data=guard
  outfile="c:\proves\secret_project\dades\guard.xlsx"
  dbms=xlsx replace;
run;

proc export
  data=victoria
  outfile="c:\proves\secret_project\dades\victoria.xlsx"
  dbms=xlsx replace;
run;

proc export
  data=comp_total_any
  outfile="c:\proves\secret_project\dades\comp_total_any.xlsx"
  dbms=xlsx replace;
run;

proc export
  data=comp_total_equip
  outfile="c:\proves\secret_project\dades\comp_total_equip.txt"
  dbms=tab replace;
run;
proc export
  data=comp_total_cluster
  outfile="c:\proves\secret_project\dades\comp_total_cluster.xlsx"
  dbms=xlsx replace;
run;

proc export
  data=test_mitjanes

```

```

        outfile="c:\proves\secret_project\dades\test_mitjanes.xlsx"
        dbms=xlsx replace;
run;

proc export
    data=test_mitjanes_prova
    outfile="c:\proves\secret_project\dades\test_mitjanes_prova.xlsx"
    dbms=xlsx replace;
run;

proc export
    data=comp_tirs_lliuers
    outfile="c:\proves\secret_project\dades\comp_tirs_lliuere.xlsx"
    dbms=xlsx replace;
run;

proc export
    data=posicio
    outfile="c:\proves\secret_project\dades\posicio.xlsx"
    dbms=xlsx replace;
run;

proc export
    data=definitiu
    outfile="c:\proves\secret_project\dades\definitiu.xlsx"
    dbms=xlsx replace;
run;

proc export
    data=definitiu_cluster
    outfile="c:\proves\secret_project\dades\definitiu_cluster.xlsx"
    dbms=xlsx replace;
run;

```

- **Codi R**

### ## Código R treball

#### # Instal.lar paquets necessaris pels diferents anàlisi.

```

install.packages("ggplot2")
install.packages("readxl")
install.packages("normtest") ###REALIZA 5 PRUEBAS DE NORMALIDAD###
install.packages("nortest") ###REALIZA 10 PRUEBAS DE NORMALIDAD###
install.packages("moments")
install.packages("dplyr")
install.packages("WRS2")
install.packages("lsr")
install.packages("corrplot")
install.packages("lattice")

```

#### # Paquets necessaris.

```

library(readxl)
library(ggplot2)

```

```

library(normtest)
library(nortest)
library(moments)
library(dplyr)
library(WRS2)
library("lsr")
library("corrplot")
library("lattice")

```

**# S'estableix un directori de treball.**

```
setwd("c:/proves/secret_project/dades")
```

**# Lectura de les diferents taules excel que s'han exportat de la part del SAS.**

```

regular<-read_excel("regular.xlsx")
finales<-read_excel("playof.xlsx")
normalitat<-read_excel("prova_normalitat.xlsx")
comp_total_any<-read_excel("comp_total_any.xlsx")
comp_total_cluster<-read_excel("comp_total_cluster.xlsx")
test_mitjanes<-read_excel("test_mitjanes.xlsx")
definitiu<-read_excel("definitiu.xlsx")
def_clust<-read_excel("definitiu_cluster.xlsx")

```

**# Es comprova que es tenen variables amb format que no interessin**

```
sapply(regular,class)
```

**# com que de l'exportació del sas tenim variables en string les modifiquem a numeriques.**

```

for (i in 3:length(regular)) {
  regular[[i]]<- as.numeric(regular[[i]])
  i<-i+1
}
for (i in 3:length(finales)) {
  finales[[i]]<- as.numeric(finales[[i]])
  i<-i+1
}
for (i in 4:length(normalitat)) {
  normalitat[[i]]<- as.numeric(normalitat[[i]])
  i<-i+1
}

```

**# Es converteix en format data frame ja que al importar directament del SAS ens ho llegeix en format diferent.**

```

regular<-as.data.frame(regular)
finales<-as.data.frame(finales)
normalitat<-as.data.frame(normalitat)

```

```

comp_total_any<-as.data.frame(comp_total_any)
test_mitjanes<- as.data.frame(test_mitjanes)
comp_total_cluster<-as.data.frame(comp_total_cluster)
definitiu<-as.data.frame(definitiu)
def_clust<-as.data.frame(def_clust)

# Transformació a variable numèrica de les categoriques per la realització de la modelització lineal.
regular[, "temp_num"]<-substr(regular[,1],1,4)
regular[, "temp_num"]<-as.numeric(regular[, "temp_num"])
comp_total_cluster[, "temp_num"]<-substr(comp_total_cluster[,2],1,4)
comp_total_cluster[, "temp_num"]<-as.numeric(comp_total_cluster[, "temp_num"])
comp_total_cluster[, "cluster_num"]<-substr(comp_total_cluster[,1],2,2)
comp_total_cluster[, "cluster_num"]<-as.numeric(comp_total_cluster[, "cluster_num"])
comp_total_any[,1]<-substr(comp_total_any[,1],1,4)
comp_total_any[,1]<-as.numeric(comp_total_any[,1])
def_clust[, "temp_num"]<-substr(def_clust[,2],1,4)
def_clust[, "temp_num"]<-as.numeric(def_clust[, "temp_num"])
def_clust[, "cluster_num"]<-substr(def_clust[,1],2,2)
def_clust[, "cluster_num"]<-as.numeric(def_clust[, "cluster_num"])
definitiu[, "temp_num"]<-substr(definitiu[,1],1,4)
definitiu[, "temp_num"]<-as.numeric(definitiu[, "temp_num"])
definitiu[, "cluster_num"]<-substr(definitiu[,29],2,2)
definitiu[, "cluster_num"]<-as.numeric(definitiu[, "cluster_num"])

# Comprovació de que ja es tenen les variables com es necessiten
sapply(regular,class)
sapply(finales,class)
sapply(normalitat,class)

# Gràfics de distribució normalitat de les diferents variables a analitzar.
hist(normalitat[normalitat$equip == "warriors" & normalitat$Temporada=="2003_2004", "3PA"])
hist(normalitat[normalitat$equip == "warriors" & normalitat$Temporada=="2010_2011", "3PA"])
x<-normalitat[normalitat$equip == "warriors" & normalitat$Temporada=="2017_2018", "3PA"]
x<-normalitat[normalitat$equip == "warriors" & normalitat$Temporada=="2017_2018", "T3_FG"]
x<-normalitat[normalitat$equip == "warriors" & normalitat$Temporada=="2017_2018", "MR_FGA"]
x<-normalitat[normalitat$equip == "warriors" & normalitat$Temporada=="2017_2018", "IP_FGA"]
x<-normalitat[normalitat$equip == "warriors" & normalitat$Temporada=="2017_2018", "RA_FGA"]
x<-normalitat[normalitat$equip == "warriors" & normalitat$Temporada=="2017_2018", "FGA"]
x<-normalitat[normalitat$equip == "warriors" & normalitat$Temporada=="2017_2018", "FTA"]
x<-normalitat[normalitat$equip == "warriors" & normalitat$Temporada=="2017_2018", "WIN_"]

# Gràfic pel redactat sobre la normalitat.

```

```

plotn <- function(x,main="Histograma Normalitat Warriors T3_FGA 2017/2018",
                xlab="T3_FGA",ylab="Densidad") {
  min <- min(x)
  max <- max(x)
  media <- mean(x)
  dt <- sd(x)
  hist(x,freq=F,main=main,xlab=xlab,ylab=ylab)
  curve(dnorm(x,media,dt), min, max,add = T,col="blue")
}

plotn(x,main="Histograma Normalitat Warriors T3_FGA 2017/2018")#Grafico de x
hist(normalitat[normalitat$equip == "jazz" & normalitat$Temporada=="2003_2004", "3PA"])
hist(normalitat[normalitat$equip == "jazz" & normalitat$Temporada=="2010_2011", "3PA"])
hist(normalitat[normalitat$equip == "jazz" & normalitat$Temporada=="2017_2018", "3PA"])
hist(normalitat[normalitat$equip == "magic" & normalitat$Temporada=="2003_2004", "3PA"])
hist(normalitat[normalitat$equip == "magic" & normalitat$Temporada=="2010_2011", "3PA"])
hist(normalitat[normalitat$equip == "magic" & normalitat$Temporada=="2017_2018", "3PA"])

# Test de Normalitat escollits.

# Prova de Shapiro-Francia
sf.test(normalitat[normalitat$equip == "warriors" & normalitat$Temporada=="2003_2004", "3PA"])
sf.test(normalitat[normalitat$equip == "warriors" & normalitat$Temporada=="2010_2011", "3PA"])
sf.test(normalitat[normalitat$equip == "warriors" & normalitat$Temporada=="2017_2018", "3PA"])
sf.test(normalitat[normalitat$equip == "jazz" & normalitat$Temporada=="2003_2004", "3PA"])
sf.test(normalitat[normalitat$equip == "jazz" & normalitat$Temporada=="2010_2011", "3PA"])
sf.test(normalitat[normalitat$equip == "jazz" & normalitat$Temporada=="2017_2018", "3PA"])
sf.test(normalitat[normalitat$equip == "magic" & normalitat$Temporada=="2003_2004", "3PA"])
sf.test(normalitat[normalitat$equip == "magic" & normalitat$Temporada=="2010_2011", "3PA"])
sf.test(normalitat[normalitat$equip == "magic" & normalitat$Temporada=="2017_2018", "3PA"])

# Shapiro Test
shapiro.test(normalitat[normalitat$equip == "warriors" & normalitat$Temporada=="2003_2004", "3PA"])
shapiro.test(normalitat[normalitat$equip == "warriors" & normalitat$Temporada=="2010_2011", "3PA"])
shapiro.test(normalitat[normalitat$equip == "warriors" & normalitat$Temporada=="2017_2018", "3PA"])
shapiro.test(normalitat[normalitat$equip == "warriors" & normalitat$Temporada=="2017_2018", "FGA"])
shapiro.test(normalitat[normalitat$equip == "warriors" & normalitat$Temporada=="2017_2018", "FTA"])
shapiro.test(definitiu[definitiu$TEAM == "Golden State Warriors", "WIN_"])
shapiro.test(definitiu[definitiu$TEAM == "Golden State Warriors", "MR_FGA"])
shapiro.test(definitiu[definitiu$TEAM == "Golden State Warriors", "RA_FGA"])
shapiro.test(definitiu[definitiu$TEAM == "Golden State Warriors", "IP_FGA"])
shapiro.test(definitiu[definitiu$TEAM == "Golden State Warriors", "T3_PC"])

```

### ## Gràfics necessaris pel redactat posterior.

```
ggplot(comp_total_any, aes(x=Temporada,y=T3_FGA))+ylim(c(10, 40))+geom_line(colour="blue")+ggtitle("Evolució llançament de 3 punts per Temporada")+geom_point(size=2,shape=21,fill="black",colour="blue")+theme_minimal()
```

```
ggplot(comp_total_any, aes(x=Temporada,y=T3_FGM))+geom_line(colour="blue")+ggtitle("Evolució llançament anotat de 3 punts per Temporada")+geom_point(size=2,shape=21,fill="black",colour="blue")+theme_minimal()
```

```
ggplot(comp_total_any, aes(x=Temporada,y=T3_FG))+ylim(c(10, 40))+geom_line(colour="blue")+ggtitle("Percentatge de llançaments anotats de 3 punts per Temporada")+geom_point(size=2,shape=21,fill="black",colour="blue")+theme_minimal()
```

```
ggplot(regular,aes(x=temp_num,y=T3_FGA,group=TEAM,colour=TEAM))+ggtitle("Tirs de 3 punts per equip")+geom_line()+geom_point(size=2,shape=21,fill="black")+theme_minimal()
```

```
ggplot(regular,aes(x=temp_num,y=MR_FGA,group=TEAM,colour=TEAM))+ggtitle("Tirs de mitja distància per equip")+geom_line()+geom_point(size=2,shape=21,fill="black")+theme_minimal()
```

```
ggplot(regular,aes(x=temp_num,y=RA_FGA,group=TEAM,colour=TEAM))+geom_line()+geom_point(size=2,shape=21,fill="black")+theme_minimal()
```

```
ggplot(regular,aes(x=temp_num,y=IP_FGA,group=TEAM,colour=TEAM))+geom_line()+geom_point(size=2,shape=21,fill="black")+theme_minimal()
```

```
ggplot(comp_total_cluster,aes(x=temp_num,y=T3_FGA,group=cluster,colour=cluster))+ggtitle("Tirs de 3 punts per cluster")+geom_line()+geom_point(size=2,shape=21,fill="black")+theme_minimal()
```

```
ggplot(comp_total_cluster,aes(x=temp_num,y=MR_FGA,group=cluster,colour=cluster))+ggtitle("Tirs de mitja distància punts per cluster")+geom_line()+geom_point(size=2,shape=21,fill="black")+theme_minimal()
```

```
ggplot(comp_total_cluster,aes(x=temp_num,y=RA_FGA,group=cluster,colour=cluster))+ggtitle("Tirs de restricted area per cluster")+geom_line()+geom_point(size=2,shape=21,fill="black")+theme_minimal()
```

```
ggplot(comp_total_cluster,aes(x=temp_num,y=IP_FGA,group=cluster,colour=cluster))+ggtitle("Tirs de in the paint per cluster")+geom_line()+geom_point(size=2,shape=21,fill="black")+theme_minimal()
```

```
ggplot(regular[regular$TEAM=="Houston Rockets",],aes(x=Temporada,y=T3_FGA,group=TEAM,colour=TEAM))+geom_line()+geom_point(size=2,shape=21,fill="black")+theme_minimal()
```

```
ggplot(regular[regular$TEAM=="Houston Rockets",],aes(x=Temporada,y=MR_FGA,group=TEAM,colour=TEAM))+geom_line()+geom_point(size=2,shape=21,fill="black")+theme_minimal()
```

```
ggplot(regular[regular$TEAM=="Houston Rockets",],aes(x=Temporada,y=RC_FGA,group=TEAM,colour=TEAM))+geom_line()+geom_point(size=2,shape=21,fill="black")+theme_minimal()
```

```
ggplot(regular[regular$TEAM=="Houston Rockets",],aes(x=Temporada,y=RA_FGA,group=TEAM,colour=TEAM))+geom_line()+geom_point(size=2,shape=21,fill="black")+theme_minimal()
```

```
ggplot(regular[regular$TEAM=="Houston Rockets",],aes(x=Temporada,y=IP_FGA,group=TEAM,colour=TEAM))+geom_line()+geom_point(size=2,shape=21,fill="black")+theme_minimal()
```

### # Apartat extra d'estudi dels Tirs lliures.

```
tirs_lliuers<-read_excel("comp_tirs_lliuere.xlsx")
```

```
tirs_lliuers<-as.data.frame(tirs_lliuers)
```

```
tirs_lliuers[,"temp_num"]<-substr(tirs_lliuers[,1],1,4)
```

```
tirs_lliuers[,"temp_num"]<-as.numeric(tirs_lliuers[,"temp_num"])
```

```
ggplot(tirs_lliuers,aes(x=temp_num,y=FTA))+ggtitle("Tirs lliures intentats per temporada")+geom_line()+geom_point(size=2,shape=21,fill="red")+theme_minimal()
```

```
ggplot(tirs_lliuers,aes(x=temp_num,y=FGA))+ggtitle("Tirs totals intentats per temporada")+ylim(c(70, 100))+geom_line()+geom_point(size=2,shape=21,fill="red")+theme_minimal()
```

### # Estudi de la posició del jugador en el canvi de llançament de 3 punts.

```
posicio<-read_excel("posicio.xlsx")
posicio<-as.data.frame(posicio)
posicio[, "temp_num"]<-substr(posicio[,1],1,4)
posicio[, "temp_num"]<-as.numeric(posicio[, "temp_num"])

ggplot(posicio,aes(x=temp_num,y=T3_FGA,group=posicio,colour=posicio))+ggtitle("Tirs intentats 3 per
posicio")+geom_line()+geom_point(size=2,shape=21,fill="black")+theme_minimal()

ggplot(posicio,aes(x=temp_num,y=MR_FGA,group=posicio,colour=posicio))+ggtitle("Tirs intentats mitja
distancia")+geom_line()+geom_point(size=2,shape=21,fill="black")+theme_minimal()

ggplot(posicio,aes(x=temp_num,y=IP_FGA,group=posicio,colour=posicio))+ggtitle("Tirs intentats aprop de la
zona")+geom_line()+geom_point(size=2,shape=21,fill="black")+theme_minimal()

ggplot(posicio,aes(x=temp_num,y=RA_FGA,group=posicio,colour=posicio))+ggtitle("Tirs intentats dintre de la
zona")+geom_line()+geom_point(size=2,shape=21,fill="black")+theme_minimal()
```

### ## Clusterització dels equips segons llançament de 3 punts.

#### # Lectura del fitxer.

```
equips<-read.table("comp_total_equip.txt",head=TRUE,sep = "\t")
View(equips)
head(equips)
summary(equips)
```

#### # Es treu la variable equips per poder fer els clusters.

```
clus<-equips[,-1]
View(clus)
```

#### # Calculem la distancia entre grups segons el # de grups.

```
set.seed(127)
bss <- kmeans(clus,centers=1)$betweenss
for (i in 2:20) bss[i] <- kmeans(clus,centers=i)$betweenss
plot(1:20, bss, type="l", xlab="Número de grupos",ylab="Sumas de cuadrados entre grupos", main="distància entre
grups")
```

#### # Calculem la distancia dins del grup segons el # de grups.

```
bss2 <- kmeans(clus,centers=1)$tot.withinss
for (i in 2:20) bss2[i] <- kmeans(clus,centers=i)$tot.withinss
plot(1:20, bss2, type="l", xlab="Número de grupos",ylab="Sumas de cuadrados intragrupos", main="distància intra
grup")
```

#### # Un cop decidit el nombre de grups fem el calcul i s'exporta.

```
cl <- kmeans(clus, 3)
writeClipboard(as.character(cl$cluster))
```

#### # Es mira quins són els centres de cada grup.

```
cl$centers
write.table(cl$centers,file = "centre_clusters.txt",sep="\t")
```

### ## Test de mitjanesper a cada anàlisi de l'estudi.

```
test_mitjanesper<-read_excel("test_mitjanesper_prova.xlsx")
```



```

shapiro.test(test_mitjanes[, "T3_FGA"])
tapply(test_mitjanes[, "T3_FGA"], test_mitjanes[, "Temporada"], shapiro.test)
var.test(test_mitjanes[, "T3_FGA"] ~ test_mitjanes[, "Temporada"])
t.test(test_mitjanes[, "T3_FGA"] ~ test_mitjanes[, "Temporada"], var.equal=TRUE)
tapply(test_mitjanes[, "MR_FGA"], test_mitjanes[, "Temporada"], shapiro.test)
var.test(test_mitjanes[, "MR_FGA"] ~ test_mitjanes[, "Temporada"])
t.test(test_mitjanes[, "MR_FGA"] ~ test_mitjanes[, "Temporada"], var.equal=TRUE)
tapply(test_mitjanes[, "RA_FGA"], test_mitjanes[, "Temporada"], shapiro.test)
var.test(test_mitjanes[, "RA_FGA"] ~ test_mitjanes[, "Temporada"])
t.test(test_mitjanes[, "RA_FGA"] ~ test_mitjanes[, "Temporada"], var.equal=TRUE)
tapply(test_mitjanes[, "IP_FGA"], test_mitjanes[, "Temporada"], shapiro.test)
var.test(test_mitjanes[, "IP_FGA"] ~ test_mitjanes[, "Temporada"])
t.test(test_mitjanes[, "IP_FGA"] ~ test_mitjanes[, "Temporada"], var.equal=TRUE)
tapply(test_mitjanes[, "IP_FGA"], test_mitjanes[, "Temporada"], shapiro.test)
var.test(test_mitjanes[, "IP_FGA"] ~ test_mitjanes[, "Temporada"])
t.test(test_mitjanes[, "IP_FGA"] ~ test_mitjanes[, "Temporada"], var.equal=TRUE)
tapply(definitiu[, "T3_PC"], definitiu[, "Temporada"], shapiro.test)
var.test(definitiu[definitiu$Temporada == "2001_2002" | definitiu$Temporada == "2021_2022", "T3_PC"] ~
definitiu[definitiu$Temporada == "2001_2002" | definitiu$Temporada == "2021_2022", "Temporada"])
t.test(definitiu[definitiu$Temporada == "2001_2002" | definitiu$Temporada ==
"2021_2022", "T3_PC"] ~ definitiu[definitiu$Temporada == "2001_2002" | definitiu$Temporada ==
"2021_2022", "Temporada"], var.equal=TRUE)
tapply(definitiu[, "FGA"], definitiu[, "Temporada"], shapiro.test)
var.test(definitiu[definitiu$Temporada == "2001_2002" | definitiu$Temporada == "2021_2022", "FGA"] ~
definitiu[definitiu$Temporada == "2001_2002" | definitiu$Temporada == "2021_2022", "Temporada"])
t.test(definitiu[definitiu$Temporada == "2001_2002" | definitiu$Temporada ==
"2021_2022", "FGA"] ~ definitiu[definitiu$Temporada == "2001_2002" | definitiu$Temporada ==
"2021_2022", "Temporada"], var.equal=TRUE)
tapply(definitiu[, "FTA"], definitiu[, "Temporada"], shapiro.test)
var.test(definitiu[definitiu$Temporada == "2001_2002" | definitiu$Temporada == "2021_2022", "FTA"] ~
definitiu[definitiu$Temporada == "2001_2002" | definitiu$Temporada == "2021_2022", "Temporada"])
t.test(definitiu[definitiu$Temporada == "2001_2002" | definitiu$Temporada ==
"2021_2022", "FTA"] ~ definitiu[definitiu$Temporada == "2001_2002" | definitiu$Temporada ==
"2021_2022", "Temporada"], var.equal=TRUE)

```

**# Estudi de correlació entre les diferents variables necessàries per l'estudi.**

```

cor(comp_total_any[, "T3_FGA"], comp_total_any[, "MR_FGA"])
cor.test(comp_total_any[, "T3_FGA"], comp_total_any[, "MR_FGA"])
cor.test(comp_total_any[, "T3_FGA"], comp_total_any[, "RA_FGA"])
cor.test(comp_total_any[, "T3_FGA"], comp_total_any[, "IP_FGA"])
cor.test(comp_total_any[, "T3_FGA"], comp_total_any[, "IP_FGA"])
cor.test(comp_total_any[, "T3_FGA"], comp_total_any[, "T3_FG"])
cor.test(def_clust[, "T3_FGA"], def_clust[, "WIN_"])

```

```

cor.test(definitiu[, "T3_FGA"], definitiu[, "WIN_"])
cor.test(definitiu[, "T3_FGA"], definitiu[, "REB"])
cor.test(definitiu[, "T3_FGA"], definitiu[, "FGA"])
cor.test(definitiu[, "T3_FGA"], definitiu[, "FTA"])
cor.test(def_clust[, "PTS"], def_clust[, "WIN_"])
cor.test(def_clust[, "PTS"], def_clust[, "WIN_"])

```

#### # Gràfics de correlació.

```

M <- cor(def_clust[, -c(1,2)])
corrplot(M, method = "ellipse")

M <- cor(definitiu[, -c(1,2,29)])
corrplot(M, method = "ellipse")

M <- cor(regular[, -c(1,2)])
corrplot(M, method = "ellipse")

cor.test(def_clust[, "cluster_num"], def_clust[, "WIN_"])

```

#### # Anàlisi gràfic del percentatge de victòria en relació al llançament de 3 punts.

```

ggplot(def_clust, aes(x=temp_num, y=WIN_, group=cluster, colour=cluster)) + ggtitle("Evolució del percentatge de
victòria per cluster") + geom_line() + geom_point(size=2, shape=21, fill="black") + theme_minimal()

ggplot(def_clust, aes(x=temp_num, y=PTS, group=cluster, colour=cluster)) + ggtitle("Tirs          intentats    3    per
posicio") + geom_line() + geom_point(size=2, shape=21, fill="black") + theme_minimal()

ggplot(def_clust, aes(x=temp_num, y=REB, group=cluster, colour=cluster)) + ggtitle("Tirs          intentats    3    per
posicio") + geom_line() + geom_point(size=2, shape=21, fill="black") + theme_minimal()

ggplot(def_clust, aes(x=temp_num, y=FTA, group=cluster, colour=cluster)) + ggtitle("Tirs          intentats    3    per
posicio") + geom_line() + geom_point(size=2, shape=21, fill="black") + theme_minimal()

ggplot(def_clust, aes(x=temp_num, y=FGA, group=cluster, colour=cluster)) + ggtitle("Tirs          intentats    3    per
posicio") + geom_line() + geom_point(size=2, shape=21, fill="black") + theme_minimal()

```

#### # Modelització lineal Variable Resposta T3\_FGA

```

model_1 <- lm(T3_FGA ~ temp_num, data = regular)
summary(model_1)

shapiro.test(residuals(model_1)) ## normalitat residus
plot(residuals(model_1))
abline(a=0, b=0) ## residus voltant

model_2 <- lm(T3_FGA ~ Temporada, data = comp_total_any)
summary(model_2)

shapiro.test(residuals(model_2)) ## normalitat residus
plot(residuals(model_2))
abline(a=0, b=0) ## residus voltant 0

model_3 <- lm(T3_FGA ~ cluster + temp_num, data = comp_total_cluster)
summary(model_3)

shapiro.test(residuals(model_3)) ## normalitat residus
plot(residuals(model_3))

```

```

abline(a=0, b=0)##residus voltant
model_4<-lm(T3_FGA ~ cluster_num+MR_FGA+IP_FGA+RA_FGA, data = comp_total_cluster)
summary(model_4)
shapiro.test(residuals(model_4))##normalitat residus
plot(residuals(model_4))
abline(a=0, b=0)##residus voltant 0
model_def<-lm(T3_FGA ~ cluster_num+MR_FGA+IP_FGA+RA_FGA, data = definitiu)
summary(model_def)
shapiro.test(residuals(model_def))##normalitat residus
plot(residuals(model_def))
abline(a=0, b=0)##residus voltant 0
plot(model_def)
model_5<-lm(T3_FGA ~ cluster_num+MR_FGA, data = def_clust)
summary(model_5)
shapiro.test(residuals(model_5))##normalitat residus
plot(residuals(model_5))
abline(a=0, b=0)##residus voltant 0
model_6<-lm(T3_FGA ~ cluster_num+MR_FGA, data = comp_total_cluster)
summary(model_6)
shapiro.test(residuals(model_6))##normalitat residus
plot(residuals(model_6))
abline(a=0, b=0)##residus voltant 0
model_7<-lm(T3_FGA ~ temp_num+cluster_num, data = comp_total_cluster)
summary(model_7)
shapiro.test(residuals(model_7))##normalitat residus
plot(residuals(model_7))
abline(a=0, b=0)##residus voltant 0

model_8<-lm(T3_FGA ~ temp_num+posicio_num+MR_FGA, data = posicio)
summary(model_8)
shapiro.test(residuals(model_8))##normalitat residus
plot(residuals(model_8))
abline(a=0, b=0)##residus voltant 0
# Modelització lineal amb variable resposta percentatge de victòria.
model_1<-lm(WIN_ ~ PTS + T3_FGA + MR_FGA, data = def_clust)
summary(model_1)
shapiro.test(residuals(model_1))##normalitat residus
plot(residuals(model_1))

```

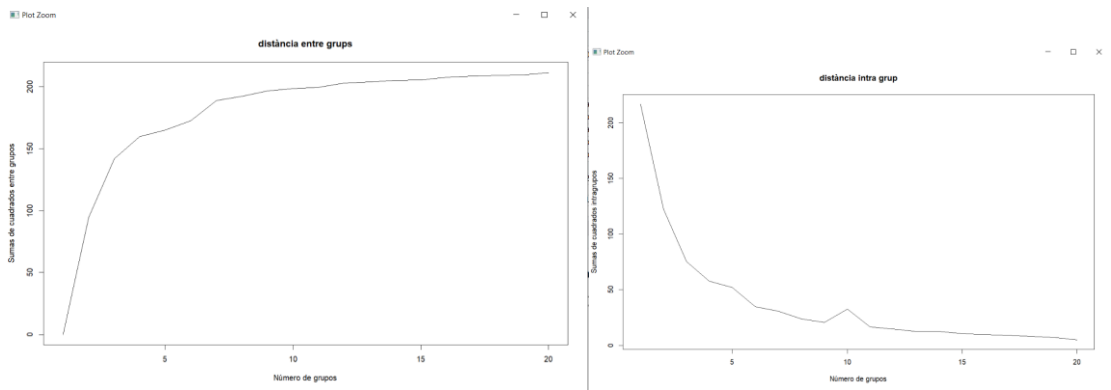
```

abline(a=0, b=0)##residus voltant 0
model_2<-lm(WIN_ ~ FGA + REB + AST + PTS + T3_FGA + MR_FGA, data = def_clust)
summary(model_2)
shapiro.test(residuals(model_2))##normalitat residus
plot(residuals(model_2))
abline(a=0, b=0)##residus voltant 0
model_3<-lm(WIN_ ~ cluster_num + FGA + REB + AST + PTS + T3_FGA + MR_FGA, data = def_clust)
summary(model_3)
shapiro.test(residuals(model_3))##normalitat residus
plot(residuals(model_3))
abline(a=0, b=0)##residus voltant 0
model_def_clust<-lm(WIN_ ~ temp_num + cluster_num + FGA + REB + AST + PTS + T3_FGA + MR_FGA, data =
def_clust)
summary(model_def_clust)
shapiro.test(residuals(model_def_clust))##normalitat residus
plot(residuals(model_def_clust))
abline(a=0, b=0)##residus voltant 0
model_def<-lm(WIN_ ~ T3_PC + FGA + PTS + REB + AST + BLK+ TOV+ STL+ FTA + temp_num, data = definitiu)
summary(model_def)
shapiro.test(residuals(model_def))##normalitat residus
plot(residuals(model_def))
abline(a=0, b=0)##residus voltant 0
# Modelització amb Anova per observar les variables categòriques.
boxplot(regular[, "T3_FGA"] ~ regular[, "TEAM"], ylab = "Tirs intentats per temporada")
fm = aov( lm(T3_FGA ~ Temporada, data = regular))
summary(fm)
intervals = TukeyHSD(fm)
intervals

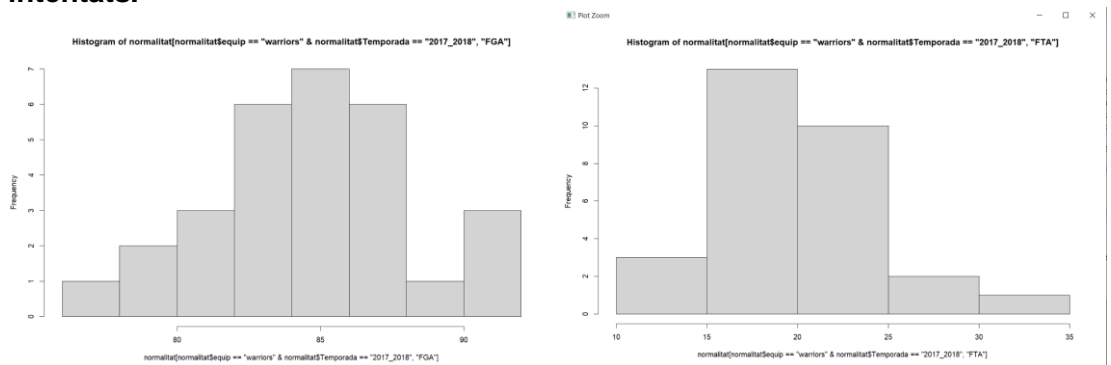
boxplot(comp_total_cluster[, "T3_FGA"] ~ comp_total_cluster[, "cluster"], ylab = "Tirs intentats per temporada")
fm = aov( lm(T3_FGA ~ cluster+Temporada, data = comp_total_cluster))
summary(fm)
intervals = TukeyHSD(fm)
intervals

```

- Gràfics per determinar el nº de clústers.



- Gràfics normalitat resta de variables llançaments totals intentats i tirs lliures intentats.



- Resultats test de normalitat de la resta de variables (excepte el llançament de 3 punts).

Shapiro-wilk normality test

```
data: normalita[normalita$equip == "warriors" & normalita$Temporada == "2017_2018", "FTA"]
W = 0.96399, p-value = 0.4103
```

Shapiro-wilk normality test

```
data: definitiu[definitiu$TEAM == "Golden State Warriors", "WIN_"]
W = 0.96965, p-value = 0.7252
```

Shapiro-wilk normality test

```
data: definitiu[definitiu$TEAM == "Golden State Warriors", "MR_FGA"]
W = 0.92373, p-value = 0.1031
```

Shapiro-wilk normality test

```
data: definitiu[definitiu$TEAM == "Golden State Warriors", "T3_PC"]
W = 0.96947, p-value = 0.7212
```

Shapiro-wilk normality test

```
data: definitiu[definitiu$TEAM == "Golden State Warriors", "IP_FGA"]
W = 0.94677, p-value = 0.2956
```

Shapiro-wilk normality test

```
data: normalitat[normalitat$equip == "warriors" & normalitat$Temporada == "2017_2018", "FTA"]
W = 0.96399, p-value = 0.4103
```

Shapiro-wilk normality test

```
data: definitiu[definitiu$TEAM == "Houston Rockets", "RA_FGA"]
W = 0.9329, p-value = 0.1573
```

- **Resultats test de mitjanes entre els llançaments de 3 punts amb la resta de variables que s'han considerat rellevants per l'estudi.**

Two Sample t-test

```
data: test_mitjanes[, "MR_FGA"] by test_mitjanes[, "Temporada"]
t = 21.705, df = 57, p-value < 2.2e-16
alternative hypothesis: true difference in means between group 2001_2002 and group 2021_2022 is not equal to 0
95 percent confidence interval:
 17.16135 20.64968
sample estimates:
mean in group 2001_2002 mean in group 2021_2022
      30.16552           11.26000
```

Two Sample t-test

```
data: test_mitjanes[, "RA_FGA"] by test_mitjanes[, "Temporada"]
t = -1.4662, df = 57, p-value = 0.1481
alternative hypothesis: true difference in means between group 2001_2002 and group 2021_2022 is not equal to 0
95 percent confidence interval:
 -2.5359356  0.3920276
sample estimates:
mean in group 2001_2002 mean in group 2021_2022
      24.54138           25.61333
```

Two Sample t-test

```
data: test_mitjanes[, "IP_FGA"] by test_mitjanes[, "Temporada"]
t = -5.9906, df = 57, p-value = 1.49e-07
alternative hypothesis: true difference in means between group 2001_2002 and group 2021_2022 is not equal to 0
95 percent confidence interval:
 -5.713275 -2.850633
sample estimates:
mean in group 2001_2002 mean in group 2021_2022
      11.84138           16.12333
```

Two Sample t-test

```
data: definitiu[definitiu$Temporada == "2001_2002" | definitiu$Temporada == "2021_2022", "T3_PC"] by definitiu[definitiu$Temporada == "2001_2002" | definitiu$Temporada == "2021_2022", "Temporada"]
t = -0.37541, df = 57, p-value = 0.7088
alternative hypothesis: true difference in means between group 2001_2002 and group 2021_2022 is not equal to 0
95 percent confidence interval:
 -1.1452425  0.7836333
sample estimates:
mean in group 2001_2002 mean in group 2021_2022
      35.17586           35.35667
```

Two Sample t-test

```
data: definitiu[definitiu$Temporada == "2001_2002" | definitiu$Temporada == "2021_2022", "FGA"] by definitiu[definitiu$Temporada == "2001_2002" | definitiu$Temporada == "2021_2022", "Temporada"]
t = -10.818, df = 57, p-value = 1.945e-15
alternative hypothesis: true difference in means between group 2001_2002 and group 2021_2022 is not equal to 0
95 percent confidence interval:
 -8.087689 -5.561277
sample estimates:
mean in group 2001_2002 mean in group 2021_2022
      81.26552           88.09000
```

```

Two Sample t-test

data: definitiu[definitiu$Temporada == "2001_2002" | definitiu$Temporada == "2021_2022", "FTA"] by definitiu[definitiu$Temporada == "2001_2002" | definitiu$Temporada == "2021_2022", "Temporada"]
t = 3.8813, df = 57, p-value = 0.0002722
alternative hypothesis: true difference in means between group 2001_2002 and group 2021_2022 is not equal to 0
95 percent confidence interval:
 0.9540737 2.9877653
sample estimates:
mean in group 2001_2002 mean in group 2021_2022
      23.82759          21.85667

```

- **Valors del test de correlació entre les variables de llançament de 3 punts i la resta de variables que s'han considerat rellevants per l'estudi.**

Pearson's product-moment correlation

```

data: comp_total_any[, "T3_FGA"] and comp_total_any[, "RA_FGA"]
t = 3.6268, df = 19, p-value = 0.001796
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.2872099 0.8394944
sample estimates:
      cor
0.6395965

```

Pearson's product-moment correlation

```

data: comp_total_any[, "T3_FGA"] and comp_total_any[, "IP_FGA"]
t = 8.3507, df = 19, p-value = 8.811e-08
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.7367592 0.9533478
sample estimates:
      cor
0.8864965

```

Pearson's product-moment correlation

```

data: comp_total_any[, "T3_FGA"] and comp_total_any[, "T3_FG"]
t = 0.67833, df = 19, p-value = 0.5057
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
-0.2976783 0.5490118
sample estimates:
      cor
0.1537683

```

Pearson's product-moment correlation

```

data: definitiu[, "T3_FGA"] and definitiu[, "WIN_"]
t = 2.9092, df = 625, p-value = 0.003753
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.0376259 0.1921478
sample estimates:
      cor
0.1155861

```

Pearson's product-moment correlation

```

data: definitiu[, "T3_FGA"] and definitiu[, "REB"]
t = 14.212, df = 625, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.4326593 0.5511882
sample estimates:
      cor
0.4942171

```

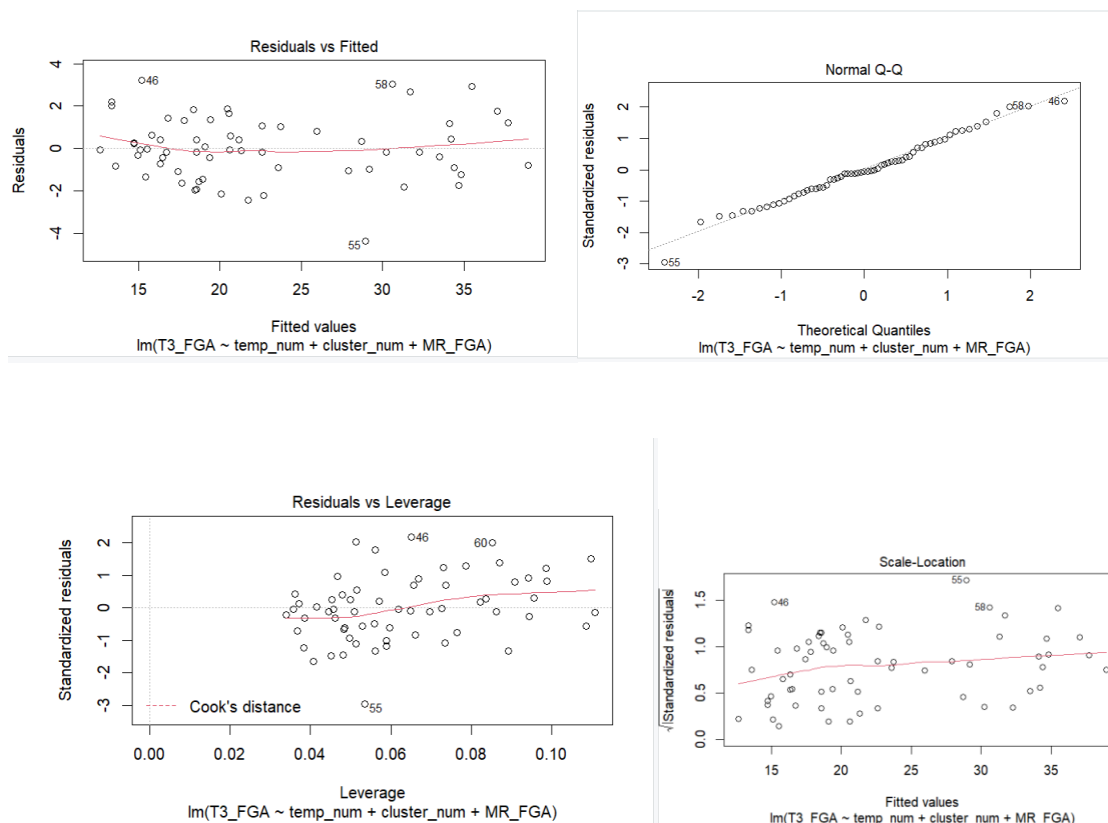
Pearson's product-moment correlation

```
data: definitiu[, "T3_FGA"] and definitiu[, "FGA"]
t = 28.463, df = 625, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.7151050 0.7835409
sample estimates:
      cor
0.751336
```

Pearson's product-moment correlation

```
data: definitiu[, "T3_FGA"] and definitiu[, "FTA"]
t = -9.0878, df = 625, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
-0.4089996 -0.2705773
sample estimates:
      cor
-0.3416399
```

- Gràfics supòsits dels residus de la modelització T3\_FGA.





- Gràfics supòsits dels residus de la modelització WIN\_.

