Pràctica 2

Apartat 1

Farem servir el dataset obtingut en la Pràctica 1 de l'assignatura, ATP_jugadors_guanys_2017 , en el qual hi trobem la llista dels primers 1000 jugadors segons els rànquing individual de l'ATP, l'Associació de Tennistes Professional, incloent els resultats esportius globals de la temporada així com els guanys obtinguts de la participació en tornejos el 2017.

L'interès en l'elaboració d'aquest dataset, o la pregunta a la qual pretén respondre, és l'anàlisi de fins a quin punt es guanyen la vida els tennistes professionals i semi-professionals. La pregunta va més encaminada a analitzar quin rendiment extreuen del tennis aquells esportistes de nivell mitjà, dins de la professionalitat, que no pas a saber el milionaris que són els millors tennistes del món.

La balança entre nivell esportiu, ingressos i renúncies és el que penso que pot tenir interès de cara a desenvolupar-ne un estudi. Per exemple, saber a quin nivell mínim s'hauria d'arribar perquè valgués la pena fer l'esforç d'intentar ser esportista d'elit en l'esport del tennis, assumint el que això podria suposar, com haver de renunciar als estudis.

Apartat 2

2.1

En el dataset original tenim els següents atributs:

Atribut	Descripció
RANKING	Número de rànquing ATP del jugador.
PAIS	Abreviatura del país del jugador.
JUGADOR	Nom del jugador.
EDAT	EDAT
PUNTS	Punts ATP del jugador.
TORNEJOS	Tornejos jugats aquesta temporada.
WL	Partits guanyats (W) - partits perduts (L) de categoria ATP World Tour, aquesta temporada pel jugador .
TITOLS	Tornejos guanyats aquesta temporada en categoria ATP World Tour.
RANKING_DOB	El rànquing de la categoria de dobles del jugador.
GUANYS_IND	Guanys en dollars aconseguits en tornejos individuals aquesta temporada.
GUANYS_DOB	Guanys en dollars aconseguits en tornejos de dobles aquesta temporada.
GUANYS_TOT	Suma de guanys individuals i dobles per jugador.

Els camps més rellevants per respondre a la pregunta que ens interessa són RANKING i GUANYS_IND. Però els altres camps ens poden ser igualment d'ajuda per a completar l'estudi. O sigui, amb aquests dos camps podem respondre a la pregunta, però si ho completem amb més camps podrem donar una resposta més completa i interessant. Per exemple, el camp WL ens determina els jugadors que han jugat tornejos a categoria ATP World Tour, la màxima categoria professional, per tant ens pot aportar una informació rellevant que ens creï subconjunts de jugadors per nivell de tornejos on participen. Per tant, tota dada ens pot aportar informació i intentarem preservar-la.

Per contra, camps com el nom del jugador (JUGADOR) no ens aporten res, ja que no volem saber què guanya un jugador particular sinó els nivells d'ingressos per nivell competitiu.

Mostrem la taula d'atributs segons el seu grau d'interès per a l'estudi (alt-mitjà-baix-nul):

Atribut	Interès
RANKING	Alt
PAIS	Anecdòtic
JUGADOR	Nul
EDAT	Baix
PUNTS	Nul – Informació totalment correlacionada amb el rànquing
TORNEJOS	Mitjà
WL	Mitjà
TITOLS	Mitjà
RANKING_DOB	No el considerarem
GUANYS_IND	Alt
GUANYS_DOB	No el considerarem
GUANYS_TOT	No el considerarem

Per interès Alt entenem que són els atributs imprescindibles per a realitzar l'estudi. Per interès mitjà entenem aquells atributs que tot i que poden aportar informació rellevant per a l'estudi, aquest es podria realitzar sense ells, encara que perdent molt poder explicatiu. El nivell Nul no ens aporta cap informació rellevant per a portar a terme concretament aquest estudi, que no vol dir que no siguin rellevants per a d'altres estudis possibles sobre aquestes mateixes dades.

En el cas del camp EDAT, li atorguem un interès baix ja que el nivell competitiu del qual en depenen directament els ingressos no depèn de l'edat del jugador, però sí que es pot estudiar si hi ha alguna mena de correlació entre edat i rànquing, pel qual més enllà dels millors tennistes, podríem donar alguna explicació al voltant d'aquest camp.

Escollim fer l'estudi per guanys individuals per seguir la lògica de les dades i no considerar ingressos que no sabem d'on surten, per això no considerem la informació del rànquing de dobles i els ingressos provinent d'aquesta disciplina. A partir d'ara, doncs, ens referirem als guanys individuals com a guanys i especificarem de quins es tracten quan no ens referim a aquests.

2.2

Les dades no contenen ni nuls ni elements en blanc. El que sí contenen són zeros. Mirem aquesta informació a través de la següent captura de pantalla:

```
> sapply(tennistes, function(x)(sum(is.na(x)))) # NA counts
    RANKING        PAIS     JUGADOR        EDAT       PUNTS    TORNEJOS          WL      TITOLS
          0           0           0           0           0           0           0           0
RANKING_DOB  GUANYS_IND  GUANYS_DOB  GUANYS_TOT
          0           0           0           0

> sapply(tennistes, function(x)(sum(x==0))) # Zero counts
    RANKING        PAIS     JUGADOR        EDAT       PUNTS    TORNEJOS          WL      TITOLS
          0           0           0           0           0           0           0         965
RANKING_DOB  GUANYS_IND  GUANYS_DOB  GUANYS_TOT
         86           3          14           2

> # Observacions amb elements en blanc

> sapply(tennistes, function(x)(sum(x==""))) # Blancs
    RANKING        PAIS     JUGADOR        EDAT       PUNTS    TORNEJOS          WL      TITOLS
          0           1           0           0           0           0           0           0
RANKING_DOB  GUANYS_IND  GUANYS_DOB  GUANYS_TOT
          0           0           0           0
```

Els atributs amb informació a zero no representen falta d'informació, sinó que contenen informació rellevant a través del número zero. Expliquem a continuació cada cas:

TITOLS: simplement el valor zero ens informa que aquell jugador no ha guanyat cap torneig a categoria ATP World Tour.

RANKING_DOB: aquest camp a zero ens indica que el jugador no ha jugat partits en categoria de dobles, per tant no ha puntuat en aquest rànquing.

GUANYS_IND: sense guanys procedents de categoria individual.

GUANYS_DOB: sense guanys procedents de categoria de dobles.

W-L: la majoria de tennistes tenen aquest camp a 0-0, això no vol dir que no hagin jugat partits, simplement vol dir que no han jugat partits individuals en la categoria màxima, ATP World Tour. Per tant, aquesta informació a zero és compatible amb haver obtingut guanys, ja que seran procedents de tornejos fora de la màxima categoria.

La problemàtica ens ve pels valors extrems, ja que ens condicionen molt les dades. Fent el càlcul de valors extrems segons el criteri de 1.5 vegades el rang interquartílic per sobre del tercer percentil, obtenim 161 observacions amb guanys individuals fora de la normalitat.

Si comparem les dades de guanys individuals pel grup de jugadors que formen part d'aquests 161 outliers amb els jugadors "normals", veiem la clara diferència:

Tennistes normals:	Tennistes amb guanys extrems:
GUANYS_IND Min. : 0 1st Qu.: 5215 Median : 9644 Mean : 21396 3rd Qu.: 22517 Max. :145770	GUANYS_IND Min. : 150224 1st Qu.: 253462 Median : 461710 Mean : 883875 3rd Qu.: 1012697 Max. :12586340

La diferència és abismal i evidentment la presència d'aquests 161 valors extrems ens distorsiona molt l'anàlisi de les dades. No cal investigar gaire per veure que aquests tennistes amb guanys desproporcionats coincideixen amb les primeres posicions del rànquing mundial. I no cal dir que es tracta de valors extrems totalment correctes, a causa dels sous desproporcionats dels primers tennistes professionals.

El tractament dels valors extrems depèn fonamentalment de l'estudi que estiguem realitzant, per tant no podem seguir un criteri general de què fer amb aquests valors perquè no existeix. Amb l'objectiu de saber fins a quin punt ens pot sortir rendible fer un esforç per dedicar-nos al tennis professional, no ens interessa saber el milionaris que són els primers tennistes del món. Ja ho sabem que ho són i podem deixar documentades les xifres, però el nostre estudi està interessat a saber fins a quin nivell hem d'arribar per guanyar-nos la vida amb el tennis. Si som els millors del món, ja no ens farem aquesta pregunta. Si la pregunta que volem respondre hagués sigut quant de rics podem arribar a ser, llavors el que ens interessaria serien precisament aquests valors extrems.

Per tant, tot i que formen part de les dades reals del món del tennis professional, decidim no considerar aquests valors extrems en les nostres dades, ja que no ens aporten cap informació perquè ja sabem que cobren moltíssims diners i no ens importa quants, però ens distorsionen les dades del món "real", del tennista professional no privilegiat. El nostre escenari d'estudi per a portar a terme el nostre objectiu coincideix plenament amb el que ens queda determinat si eliminem els valors extrems de les nostres dades. Tot i així, mantindrem aquesta informació per si cal fer alguna comparació.

Apartat 3

3.1

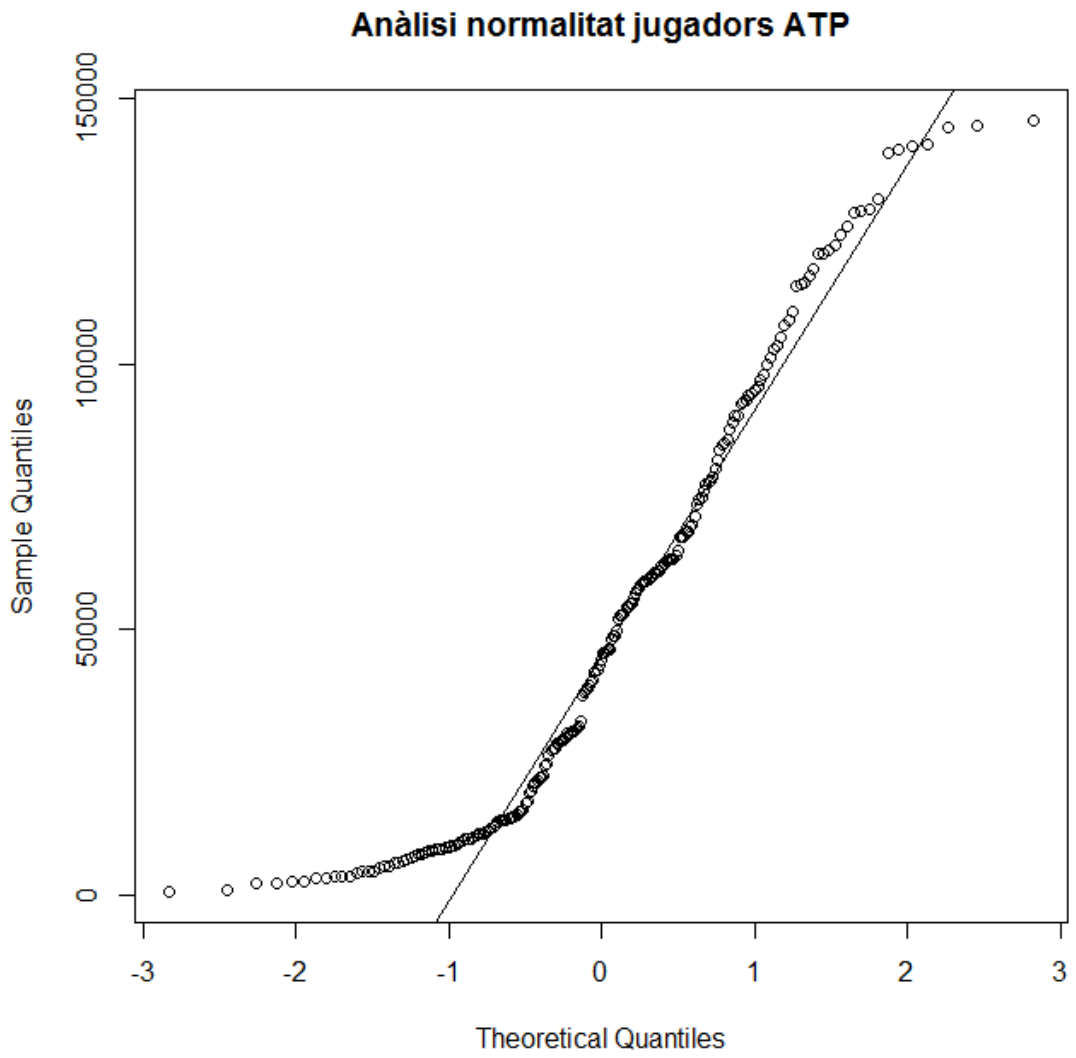
La selecció dels grups de dades que anàvem a fer queda qüestionada per la decisió anterior de no considerar els jugadors professionals que cobren fora del normal. Però el fet de considerar-los només ens distorsionava les dades, ja que ja sabem que els millors cobren molt i el que cobren queda fora del nostre interès d'estudi. Seguim, igualment, amb el que volíem fer per veure si encara té sentit. Els grups que voldríem generar són els que formarien els jugadors que han jugat algun partit a la màxima categoria del tennis, l'ATP World Tour, i els jugadors que tot i ser professionals no han jugat cap partit en aquesta categoria. Ho sabem a través dels atributs WIN LOSE, informació que hem separat procedent de l'atribut W-L. Si estan tots dos a zero, vol dir que no han jugat partits a la màxima categoria.

Mirem com ens queden aquests possibles grups, com hem dit, simplement mirant si tenen o no tenen els atributs WIN i LOSE a zero. El resultat és que tenim el grup de tennistes professionals de categoria ATP World Tour amb 212 observacions. L'altre grup NO-ATP en té 627. Com que tot i la reducció de 161 tennistes amb guanys fora del normal que formarien part del grup ATP, seguim tenint 212 observacions en aquest grup, mantenim aquesta divisió.

3.2

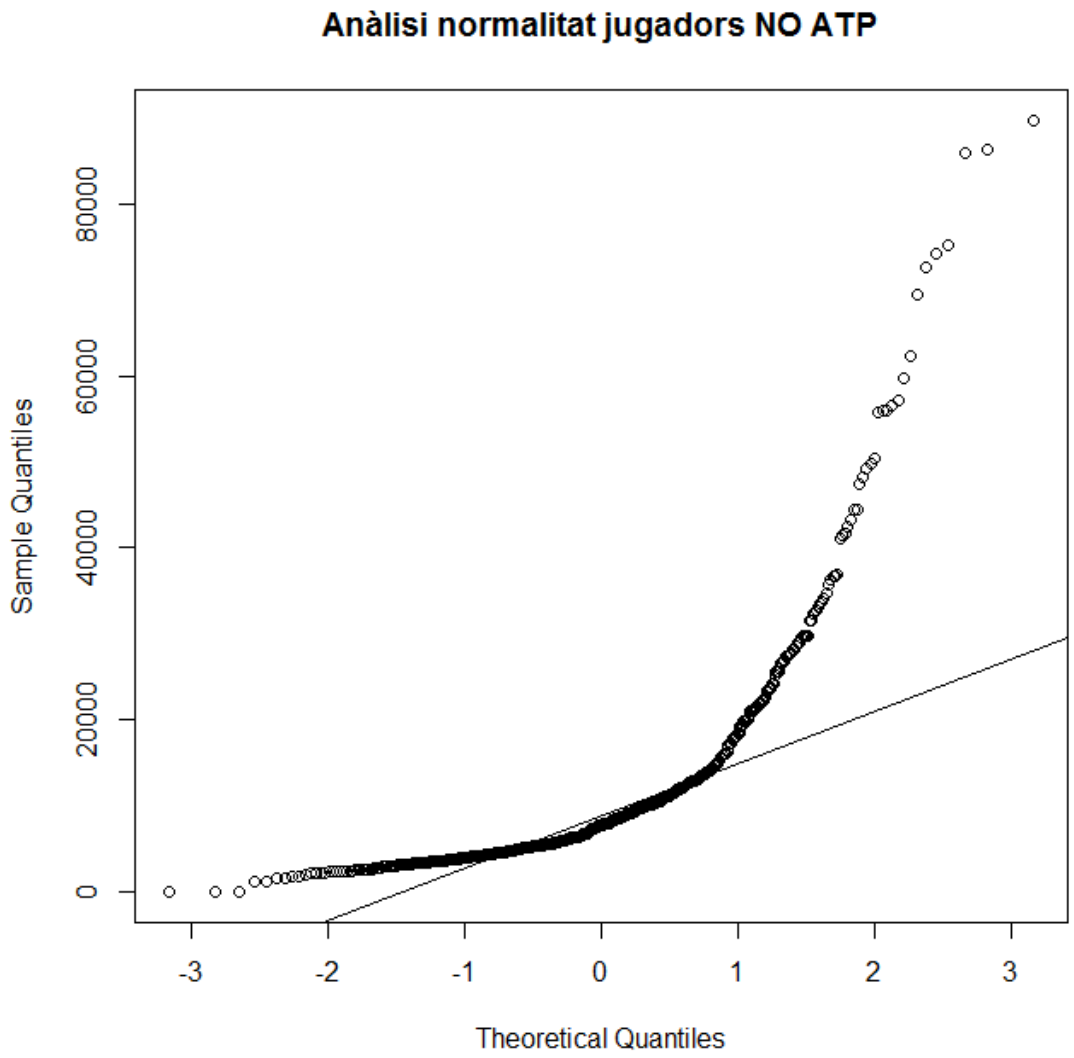
Mirem la normalitat de les dades i concloem que les dades no segueixen una distribució normal. Creiem, però, que augmentant la mostra de dades acabaríem obtenint la distribució normal, ja que tenim franges de normalitat, però que se'ns acaben escapant pels extrems. Fent testos numèrics de normalitat, que trobarem en el codi R adjunt, ens surten tots, i de molt tros, no normals.

Mostrem uns quants exemples a través de gràfiques:



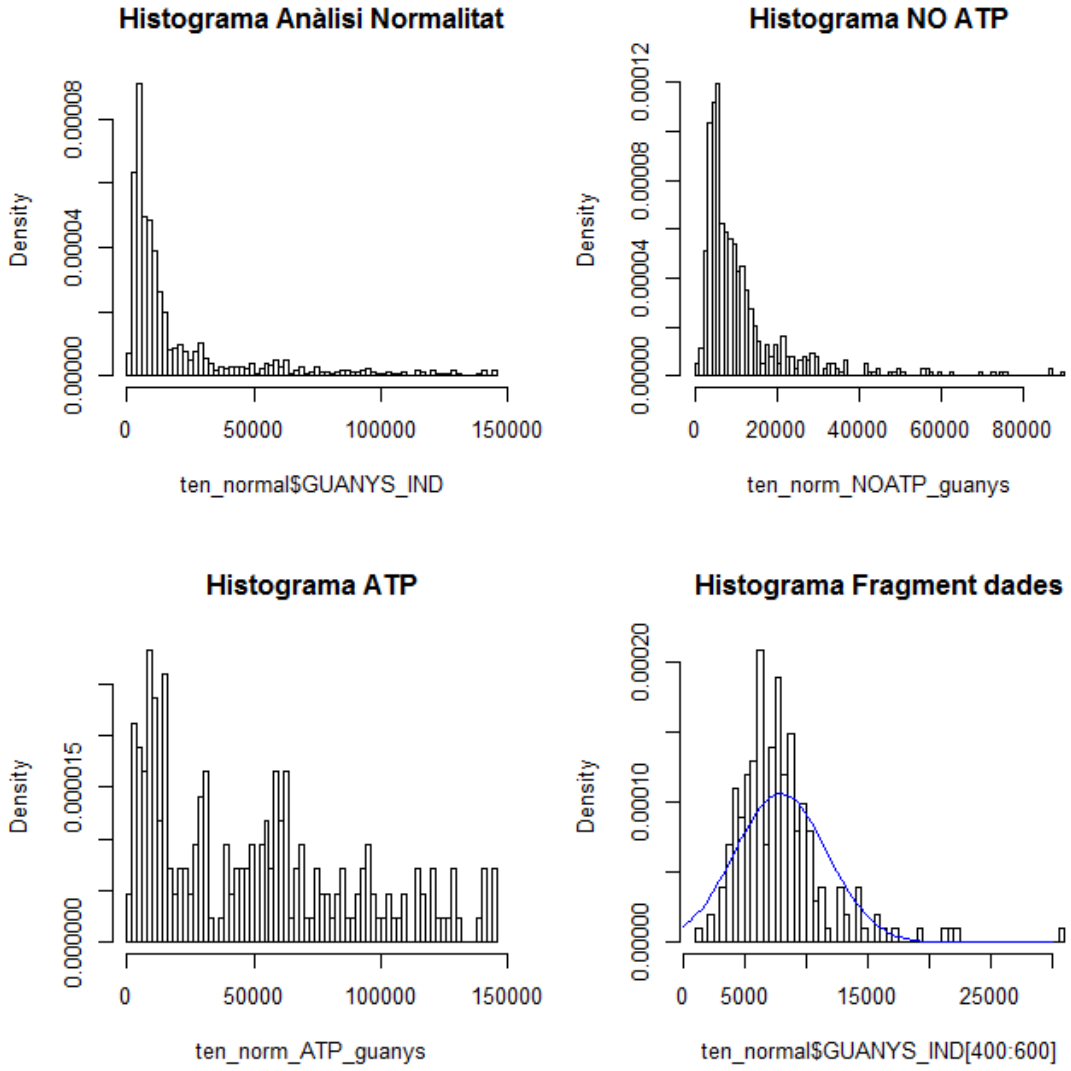
Fent l'anàlisi a través de QQ-Plot (Quantile-Quantile Plots) i en el cas del jugadors del grup ATP, sempre dins de les dades sense outliers, podem comprovar que durant almenys algunes franges segueix la distribució normal, però se'ns escapa totalment per la part inferior i superior.

En el cas del grup de jugadors No ATP:

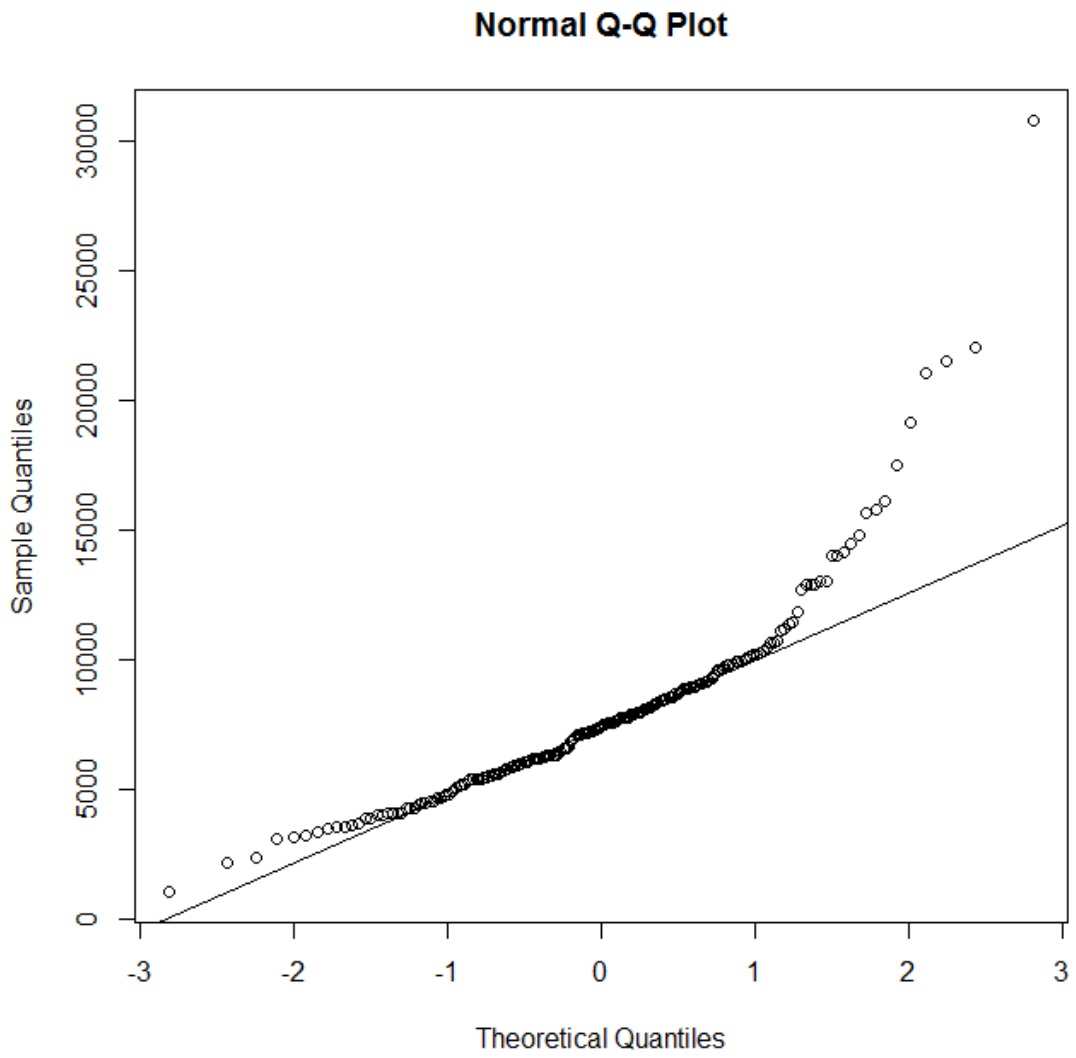


Aquí la diferència ja és abismal respecte la línia que hauríem de seguir en una distribució normal.

A continuació mostrarem diferents histogrames per analitzar la normalitat de la distribució de dades:



L'anàlisi de normalitat del cas de l'última gràfica, la del fragment, a través de QQ-Plot ens dóna el següent resultat:



Per tant, n'obtenim un resultat molt ajustat en la part central i se'ns torna a escapar pels extrems.

Pel que fa a la homogeneïtat de la variància, hem realitzat tres testos diferents, amb el resultat que tots indiquen clarament i rotundament la falta d'homogeneïtat de la variància. Els testos que hem realitzat són els de Bartlett, Levene i el test F, que també es poden trobar en el codi R adjunt.

3.3

Comencem amb un resum estadístic dels atributs que ens interessen de les dades dels tennistes sense valors extrems:

```
> summary(ten_normal[c(1,2,4,6,10,13,14,15)])
    RANKING           PAIS          EDAT          TORNEJOS       GUANYS_IND          WIN
 Min.   :  88.0   USA    : 67   Min.   :17.00   Min.   : 1.00   Min.   :     0   Min.   :0.0000
 1st Qu.: 370.5   FRA    : 62   1st Qu.:21.00   1st Qu.:14.00   1st Qu.:  5215   1st Qu.:0.0000
 Median : 581.0   ITA    : 54   Median :24.00   Median :20.00   Median :  9644   Median :0.0000
 Mean   : 579.0   ESP    : 49   Mean   :24.16   Mean   :18.84   Mean   : 21396   Mean   :0.2706
 3rd Qu.: 790.5   GER    : 47   3rd Qu.:27.00   3rd Qu.:24.00   3rd Qu.: 22517   3rd Qu.:0.0000
 Max.   :1000.0   RUS    : 37   Max.   :41.00   Max.   :45.00   Max.   :145770   Max.   :8.0000
                  (Other):523
      LOSE          ATP
 Min.   :0.0000   ATP  :212
 1st Qu.:0.0000   NOATP:627
 Median :0.0000
 Mean   :0.4529
 3rd Qu.:0.0000
 Max.   :6.0000
```

La informació de País l'hem mantinguda a nivell d'anècdota i, com podem veure, ja hem generat els atributs WIN i LOSE i ATP, que ens selecciona el grup que pertoca per cada jugador.

Introduïm també la correlació entre els guanys i tots els camps numèrics:

```
> corGuanysN
   Rànking       Edat    Tornejos         WIN        LOSE
-0.7032663  0.1471170  0.2912195  0.4748833  0.6896861
```

Sabem que la informació dels partits guanyats i perduts només reflecteix els jugats a la màxima categoria ATP, per això es dona el fet curiós que tenim més correlació de derrotes amb guanys que no pas de victòries amb guanys. Això és degut a que no considerem els millors tennistes ATP, per tant tenim més derrotes a la màxima categoria que victòries. Però igualment els millors tornejos donen molts diners només per jugar-hi, per això tenim les derrotes correlacionades amb els guanys.

Si ho comparem amb el cas dels millors jugadors del món:

```
> corGuanysO
    Rànking        Edat    Tornejos         WIN        LOSE      TITOLS
-0.45387568  0.13242170 -0.16966972  0.71927212  0.13946456  0.06487894
```

Evidentment, en condicions de normalitat, o sigui, que els jugadors tenen més diners perquè juguen sempre en els tornejos de més categoria, o sigui, sense canviar de cop la proporcionalitat dels tornejos que jugues, llavors els guanys estan més correlacionats amb les victòries que amb les derrotes. El títols no tant, ja que de tennistes que guanyin títols n'hi ha molt pocs, però pots guanyar molts diners si ets finalista a molts tornejos i no en guanyes cap. De fet, en el cas dels tennistes de nivell mitjà, precisament les dades que considerem per a l'estudi, no hem mostrat la informació de títols perquè és zero. El rànquing surt correlacionat negativament en ambdós casos, ja que com més amunt del rànquing i, per tant, amb el

número més baix, normalment guanyes més diners perquè vol dir que has guanyat més partits i conseqüentment tindràs accés a jugar tornejos de més pes econòmic.

Seguim amb més dades estadístiques del conjunt de dades sense outliers comparades per grups ATP o No ATP:

```
> edat
$ATP
     mitjana   var desv.est min    Q1 Q2 Q3 max
[1,]   24.93 18.31     4.28  17 21.75 24 28  38

$NOATP
     mitjana   var desv.est min Q1 Q2 Q3 max
[1,]   23.91 14.71     3.84  17 21 23 26  41


> rank
$ATP
     mitjana       var desv.est min     Q1    Q2     Q3 max
[1,]  402.49 51510.58   226.96  88 227.75 327.5 550.25 995

$NOATP
     mitjana       var desv.est min    Q1  Q2    Q3  max
[1,]  638.66 49121.87   221.63 134 459.5 654 828.5 1000
```

```
> guanys
$ATP
      mitjana         var desv.est min       Q1    Q2       Q3    max
[1,] 50275.68 1587136946 39838.89 774 13956.25 44808 76368.75 145770

$NOATP
      mitjana       var desv.est min     Q1   Q2    Q3   max
[1,] 11631.91 152912667 12365.79   0 4649.5 7686 12893 89736
```

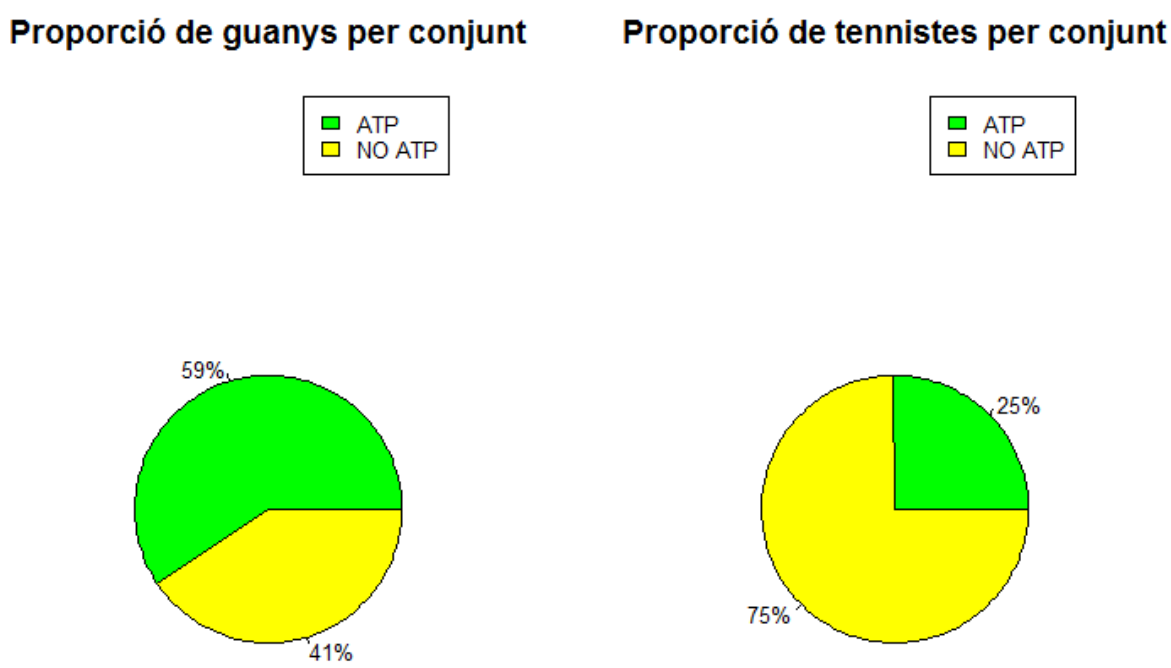
L'edat és semblant en ambdós conjunts de dades, el rànquing varia, ja que els jugadors que han participat en tornejos ATP estaran més amunt. I el més important que ens ocupa, és que els guanys també varien i molt. Tot i no tenir els tennistes que guanyen molts diners a les dades, la diferència entre jugar ATP o no és abismal. Veiem que la mitjana és d'uns 50.000 dollars i la mediana 44.808, que és un bon sou si tenim com a referència el guanyar-nos la vida. Per tant tenim que la meitat de les observacions de guanys ens servirien per a aquest objectiu. En canvi, si mirem els jugadors de fora els tornejos ATP, ens hem de situar entre el 3r quartil i el màxim per trobar un sou semblant. També és curiós que la mitjana del rànquing ATP dels jugadors que han jugat tornejos ATP se situa al número 402, amb la millor posició al 88. En

canvi els jugadors No ATP, tenen la millor posició del rànquing en la 134, realment molt alta. Després ja es normalitza la situació i el primer percentil ja cau fins a la posició 459.

Apartat 4

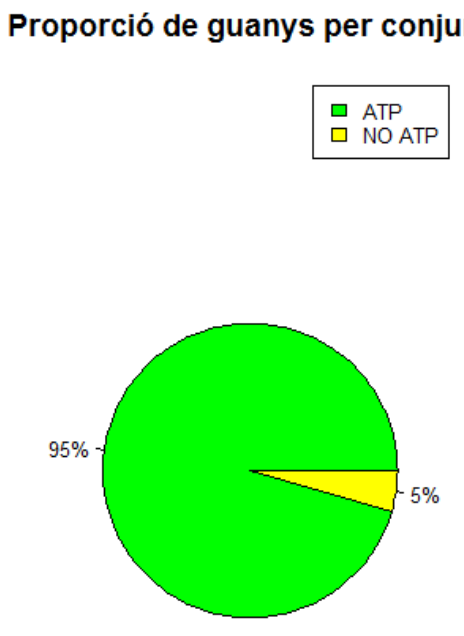
Canviem d'apartat, tot i que no deixa de ser una continuació de l'anterior, ja que és un procés continuo que ens ha de portar als resultat i conclusions finals. En l'apartat anterior ja hem anat mostrant taules de dades i gràfiques. Seguirem amb el procés de descoberta d'informació que contenen les dades, per passar a mostrar alguns resultats obtinguts.

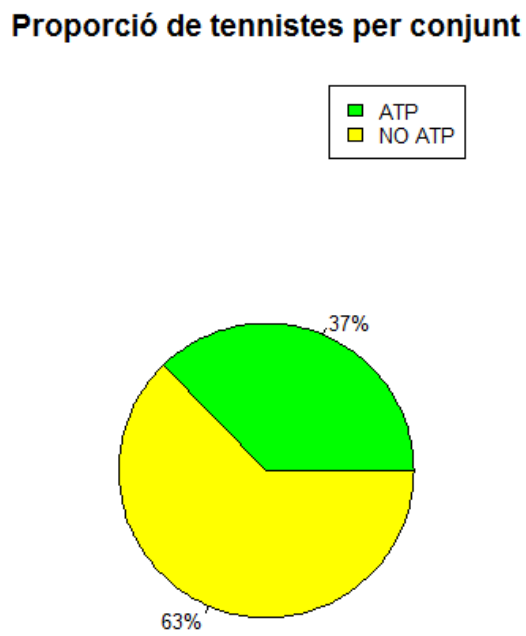
Destaquem la importància d'arribar a jugar tornejos ATP World Tour a través d'aquesta gràfica combinada:



Per tant, ja només dels tennistes sense valors extrems, o sigui, sense els que guanyen molts diners, els que juguen tornejos ATP s'emporten el 59% dels guanys amb només el 25% de tennistes.

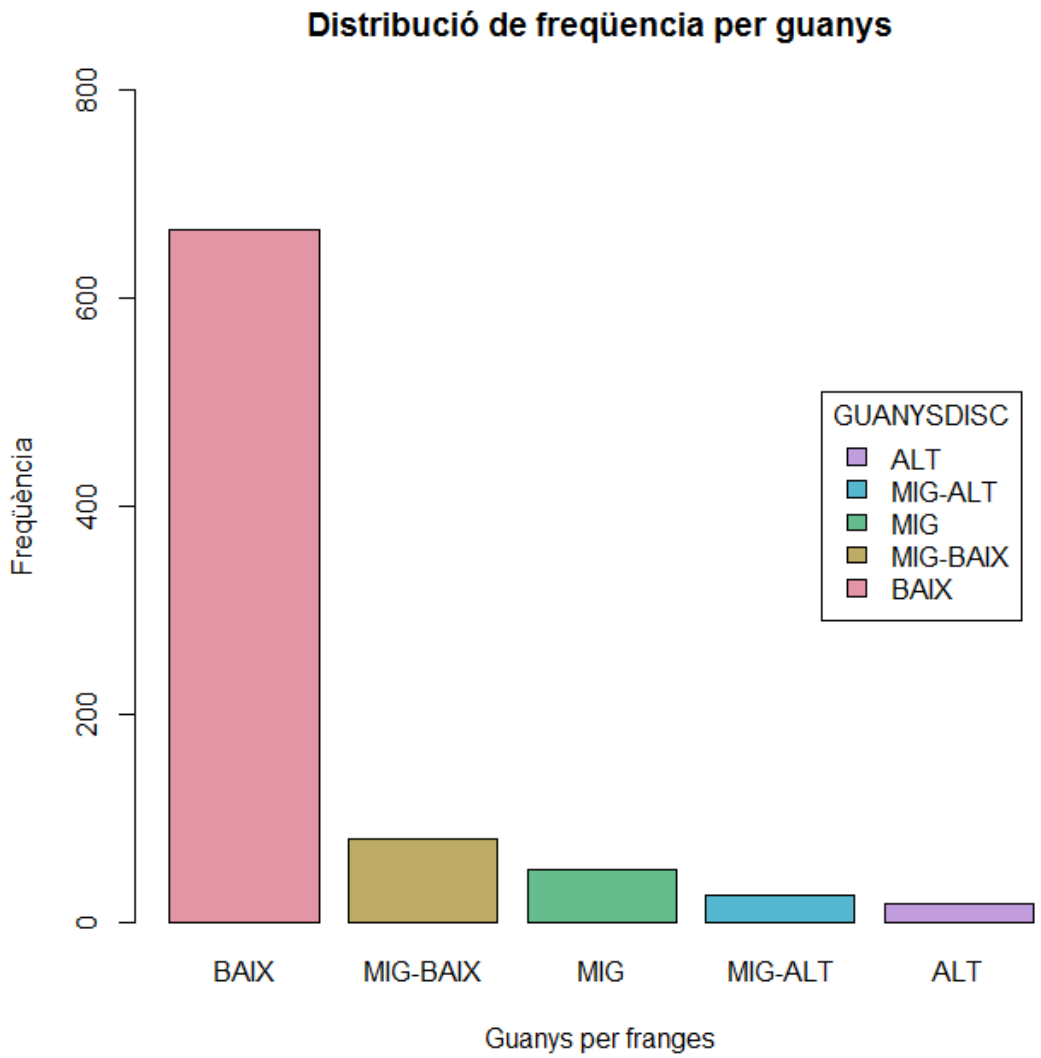
Si incloem tots els tennistes, sense eliminar els que guanyen més, la desigualtat creix:



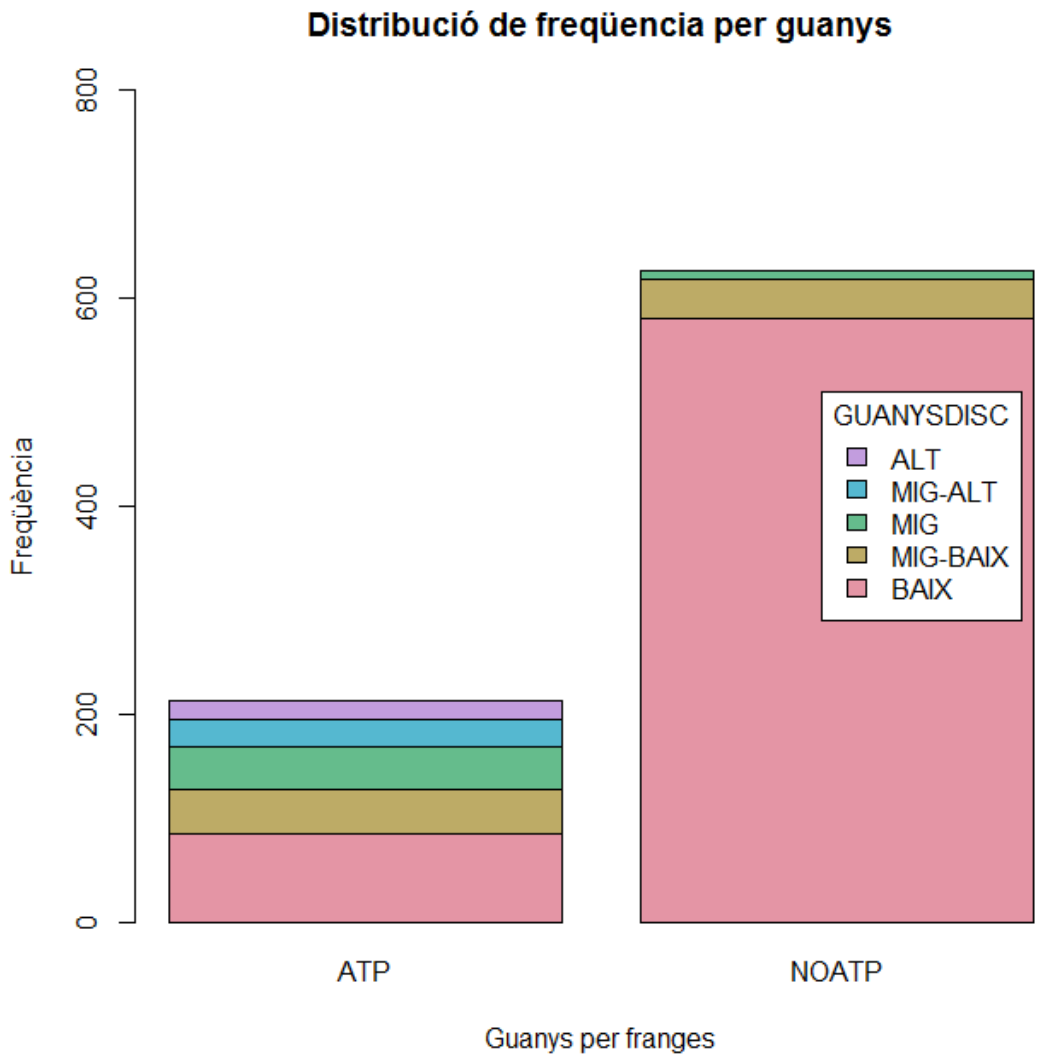


Així, els tennistes que no juguen tornejos ATP només obtenen un 5% dels guanys dels tennistes professionals, essent el 63% dels 1000 primers tennistes del món.

Seguim analitzant els guanys a través de discretitzar l'atribut de guanys individuals i veure quantes observacions tenim de cada:



La mateixa gràfica però per grup ATP o NO ATP:



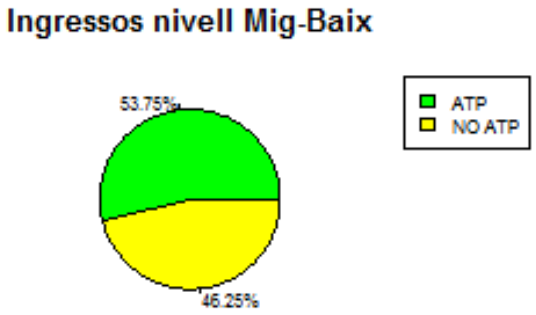
Els dos grups segueixen la mateixa lògica de tenir més observacions als nivells de guanys baixos i progressivament el nombre d'observacions van baixant a mesura que augmentem els guanys. La diferència està en l'evident desproporció entre nivells del grup de jugadors NOATP i, a més, que aquest grup en prou feines arriba a tenir observacions del interval mitjà.

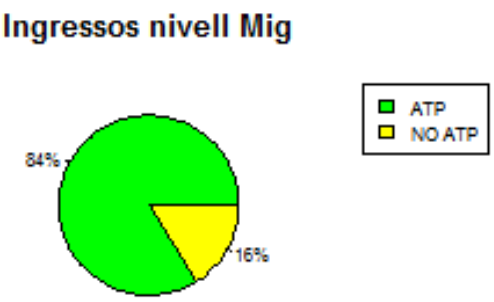
Si mirem els valors numèrics corresponents a aquestes dades, trobarem alguna informació més detallada:

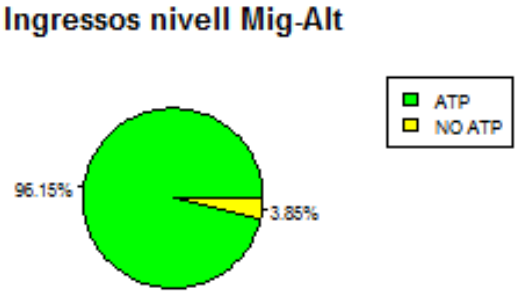
	BAIX	MIG-BAIX	MIG	MIG-ALT	ALT
meanATP	12191.202	43393.93	68826.00	100357.92	131600.28
minATP	774.000	29222.00	58513.00	87774.00	117807.00
maxATP	28802.000	57919.00	85841.00	116603.00	145770.00
meanNOATP	8831.499	40166.65	73275.38	89736.00	NA
minNOATP	0.000	29547.00	59752.00	89736.00	NA
maxNOATP	28870.000	57242.00	86389.00	89736.00	NA
countATP	84.000	43.00	42.00	25.00	18.00
countRelATP	39.620	20.28	19.81	11.79	8.49
countNOATP	581.000	37.00	8.00	1.00	0.00
countRelNOATP	92.660	5.90	1.28	0.16	0.00
countRelATPTOT	12.630	53.75	84.00	96.15	100.00
countRelNOATPTOT	87.370	46.25	16.00	3.85	0.00

Evidentment, la mitjana per rang d'ingressos és semblant en ambdós conjunts, perquè els intervals són els mateixos, la diferència està en la freqüència d'observacions en cada interval, on els jugadors NO ATP formen part del rang més baix d'ingressos en un 92,66%, per un 39,62 dels jugadors que han jugat algun partit a categoria màxima. Aquesta proporcionalitat en les observacions per nivell del cas del grup ATP i desproporcionalitat en el NO ATP, ens portarà, com veurem més endavant, a la no presència d'outliers en el primer cas i la nombrosa presència en el cas NO ATP, ja que tenim pràcticament totes les observacions a un extrem dels possibles interval.

Acabem de veure i comentar la distribució de freqüència de guanys per a cada interval de cada grup, mirem ara el mateix però de manera intergrupal, o sigui, la proporció d'observacions per grup de cada nivell d'ingressos. Obtenim les següents gràfiques:







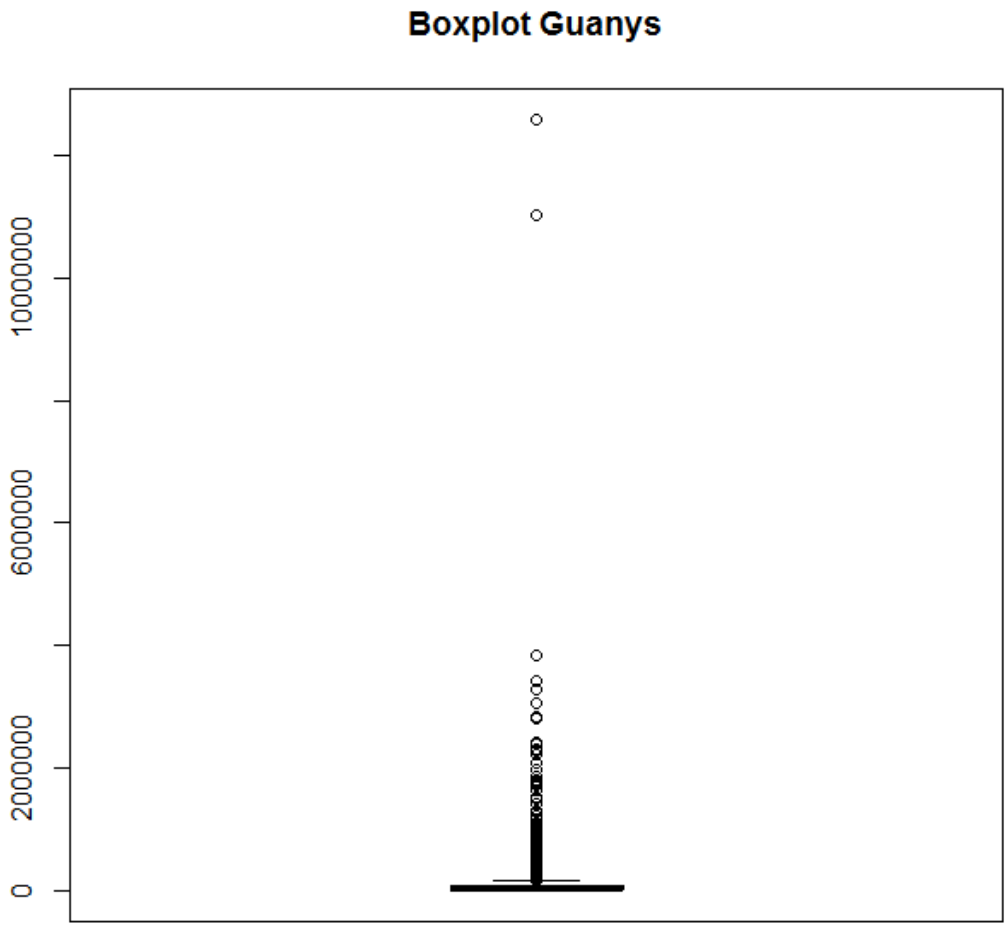
Per guanyar-nos una mica bé la vida amb el tennis professional, hauríem d'anar a parar al rang MIG-BAIX d'ingressos. Mirem a quin rànquing correspondria:

```
> rankMigBaixInfo
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  155.0   225.5   266.0   288.1   319.8   899.0

> sort(rankMigBaix)
 [1] 155 166 172 178 179 180 181 186 189 202 204 210 214 215 216 218 219 220 223 224 226 227 228
[24] 230 231 233 234 239 240 244 246 251 253 255 256 257 260 261 262 265 267 268 272 273 278 279
[47] 282 283 284 285 286 288 289 302 306 307 308 317 318 319 322 334 337 338 339 342 348 354 364
[70] 379 381 410 415 426 447 449 454 482 571 899
```

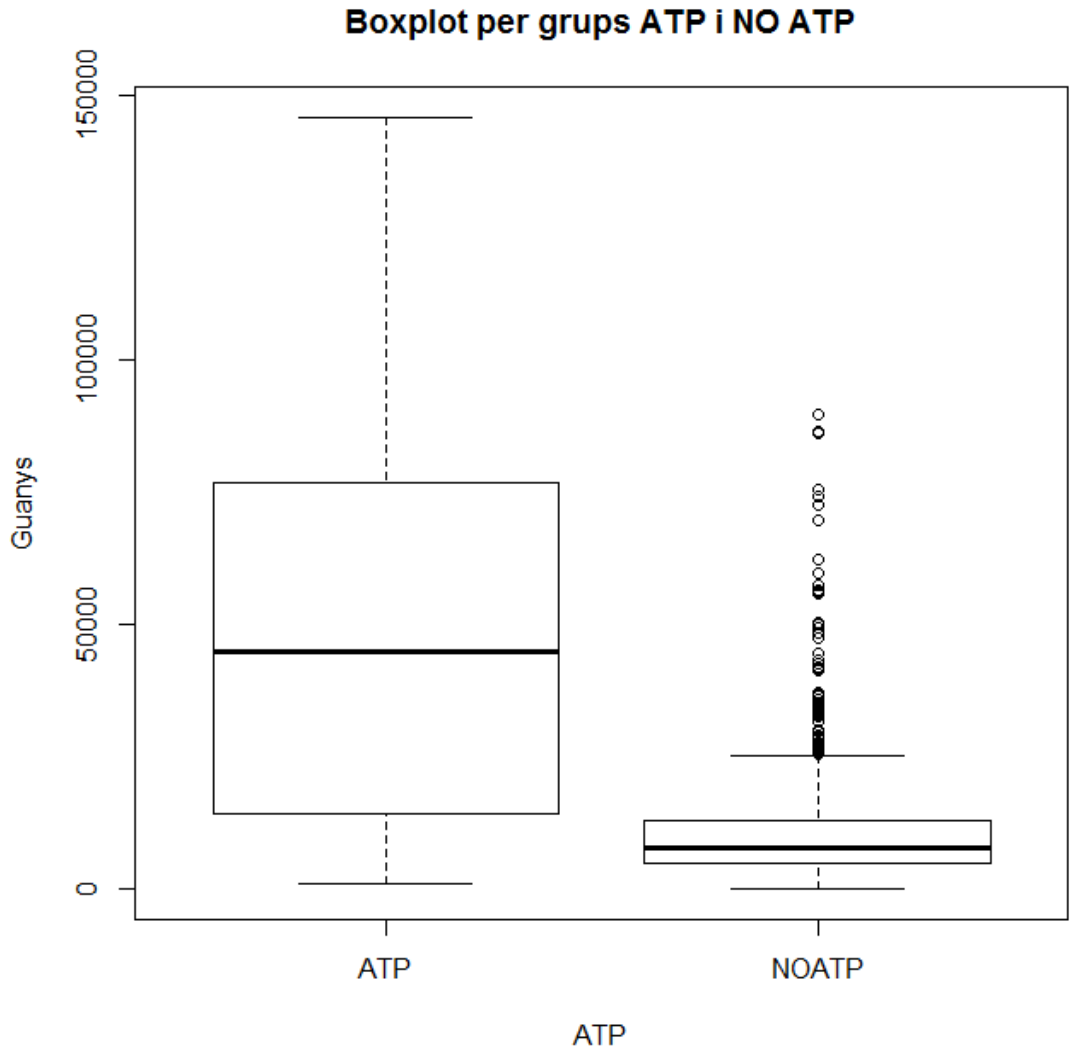
El pitjor rànquing per obtenir aquests ingressos del rang MIG-BAIX és el 899, però es tracta d'una excepció. La mitjana es situa al rànquing 288, però a partir del 454 ja es va estabilitzant dins d'aquest marge de guanys. Podem explicar els valors dels extrems, ja que segurament el jugador amb rànquing 899 està fent una bona temporada, en aquest context de jugador de nivell mitjà, i està pujant en el rànquing. Per tant, té uns ingressos que no corresponen o que encara no s'han vist reflectits en el rànquing. D'igual manera però a la inversa, segurament el jugador 155 del món té pocs ingressos perquè està fent una mala temporada, però està molt amunt en el rànquing pels bons resultat d'anys anteriors. Això explicaria els pocs ingressos que té amb aquest rànquing tan elevat.

Acabem per on havíem començat i mostrem el boxplot dels guanys dels tennistes incloent també els millors del món:



Com es pot comprovar, la presència dels outliers ens distorsionava totalment les dades, la caixa del bloxplot pràcticament ni apareix, convertida en una línia gruixuda. Si mirem els valors dels valors extrems, veiem que es tracta dels tennistes milionaris.

I, finalment, el boxplot comparatiu dels dos grups ATP i NO ATP dels tennistes que hem estudiat, per tant sense els millors i més rics del món:



No tenen res a veure, sobretot per la presència de valors extrems dins del propi grup NO ATP, ja que els valors extrems dels guanys del grup ATP són precisament els que hem deixat fora de l'anàlisi. En canvi, els valors extrems del grup NO ATP, no són valors extrems en el conjunt de les dades considerades, però sí en relació al seu grup. Això ho hem vist clarament en la seva distribució de valors, on el 92,6% forma part de la franja d'ingressos més baixa.

Apartat 5

Com a primera conclusió, hem vist com hem pogut respondre a la pregunta objectiu a través de l'anàlisi de les dades. La resposta a la pregunta és que, si algú es vol dedicar al tennis professional i guanyar-se mitjanament bé la vida pel qual valgui la pena l'esforç d'intentar-ho, el nivell mínim que hauria d'assolir està al voltant del 450 del rànquing mundial, per tenir uns ingressos a partir dels 30.000 dollars.

Per tenir una referència de la dificultat d'assolir aquest objectiu, a la Primera divisió de la lliga espanyola de futbol hi ha 482 jugadors. Però de lligues de futbol professionals almenys n'hi ha a pràcticament tots els països europeus i americans. Per tant, si ho comparem amb el tennis, tenim que les opcions se'ns redueixen moltíssim, ja que seria l'equivalent a fer-se lloc a una única lliga de futbol a nivell mundial. I si hi arribem a jugar, partiríem d'un mínim de 30.000 dollars per any. Per tant, a priori sembla que és molta competència per una promesa d'ingressos normals o prou bons en el nostre entorn, però molt baixos pel que fa l'esport i més amb aquest nivell de competència.

D'esportistes milionaris en el món del tennis, com hem vist, a partir del 150.000 dollars l'any, en tenim 161, els que hem assenyalat com a valors extrems. Llavors, si som uns privilegiats, ens podrem enfilar entre els 2,4 i 12 milions de dollars l'any, formant part del top ten del tennis mundial.

La dificultat a l'hora d'analitzar aquestes dades és la volatilitat de les mateixes. En un context on la majoria no es guanyen gaire bé la vida, però quan cobres, cobres moltíssim i molt desproporcionadament per sobre dels altres, és difícil no tenir valors extrems pràcticament sempre, perquè el que et fa passar d'un valor normal a un valor extrem quant als guanys, pot ser simplement treure el nas i jugar un simple partit en algun dels tornejos de l'ATP World Tour. Per tant, amb un fet molt puntual, en qualsevol moment podem trencar la lògica de la nostra classificació al rànquing mundial. En aquest mateix sentit, hem observat jugadors que no tenen el nivell d'ingressos que els pertocaria per la seva classificació i és degut al fet que acabem de mencionar, ja que un petit èxit et podria catapultar en ingressos i la classificació mundial tardaria una mica més a reflectir aquest èxit. Dit d'una altra manera, la classificació del rànquing mundial es comporta de forma molt més estable que els ingressos, que de seguida queden desproporcionats per petits èxits. Fins a tal punt és així, que si tenim la sort de participar en algun d'aquests tornejos de màxim nivell i millor pagats, ni tan sols ens fa falta cap victòria, ja que hem observat la curiosa correlació entre derrotes i guanys en aquests casos on tennistes de nivell mitjà treuen el nas a la màxima categoria del tennis.