

Examen Rattrapage Data Exploration 2019-2020

ING1-GI

Cette épreuve contient 2 Parties indépendantes, les questions portent principalement sur votre capacité l'analyse d'interprétation et de synthèse. Les réponses sont à rendre sous forme de document pdf de préférence, ou bien un scan ou photo des réponses manuscrites.

Partie 1 : Analyse et Interprétation de Résultats

Dans cette partie, nous vous présentons des résultats issus d'un projet effectué par des élèves ingénieurs passionnés par le Tennis. Plusieurs méthodes vues dans le cours ont été testées, nous vous présentons ci-dessous les résultats à interpréter.

Présentation du jeu de données et objectif de l'étude :

- **Jeu de données :**
 - 50 Individus : les 50 premiers joueurs mondiaux de l'année 2017.
 - 25 Variables : 22 quantitatives et 3 qualitatives.
- **Objectif :** l'intérêt principal de l'étude était l'analyse de l'influence du service en fonction des deux surfaces de jeu majeures (dur et terre battue). L'étude menée visait donc à voir si :
 - La taille du joueur constituait un élément déterminant au service
 - La surface de jeu avait une influence sur le service
 - Le gaucher était favorisé au service

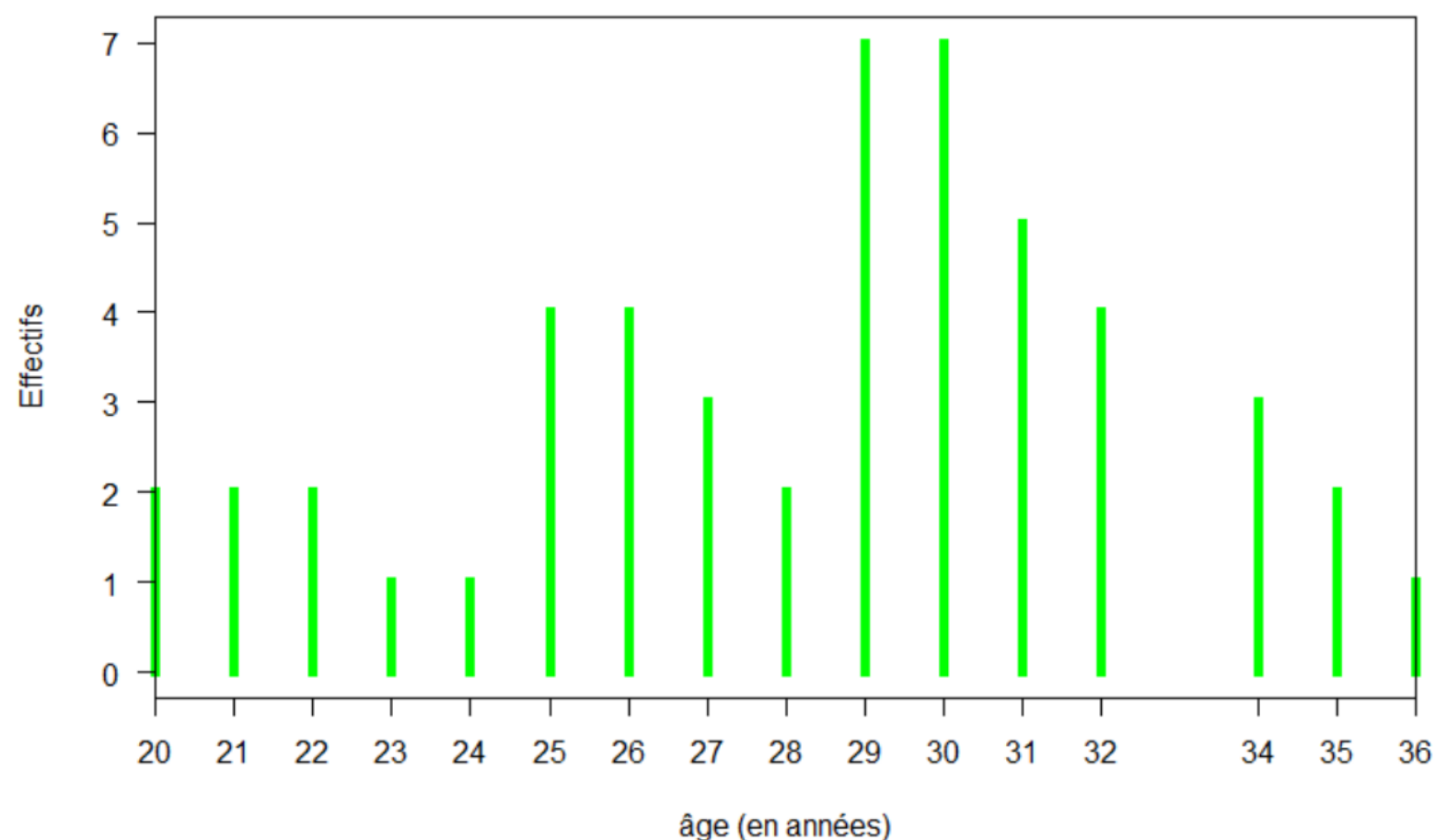
Nous vous demandons ci-dessous d'analyser et d'interpréter une série de résultats (numéroter de A à K) issus de ce projet et qui mettent en œuvre plusieurs méthodes d'exploration de données vues dans le cours.

Votre analyse et interprétation pour chaque graphique et tableau (de A à K) doit inclure les points suivants :

- 1) décrire la nature (possible) de la variable
- 2) la méthode utilisée
- 3) la constatation qu'on pourrait tirer des graphiques et des tableaux présentés
- 4) s'il y a lieu des propositions d'amélioration.

A)

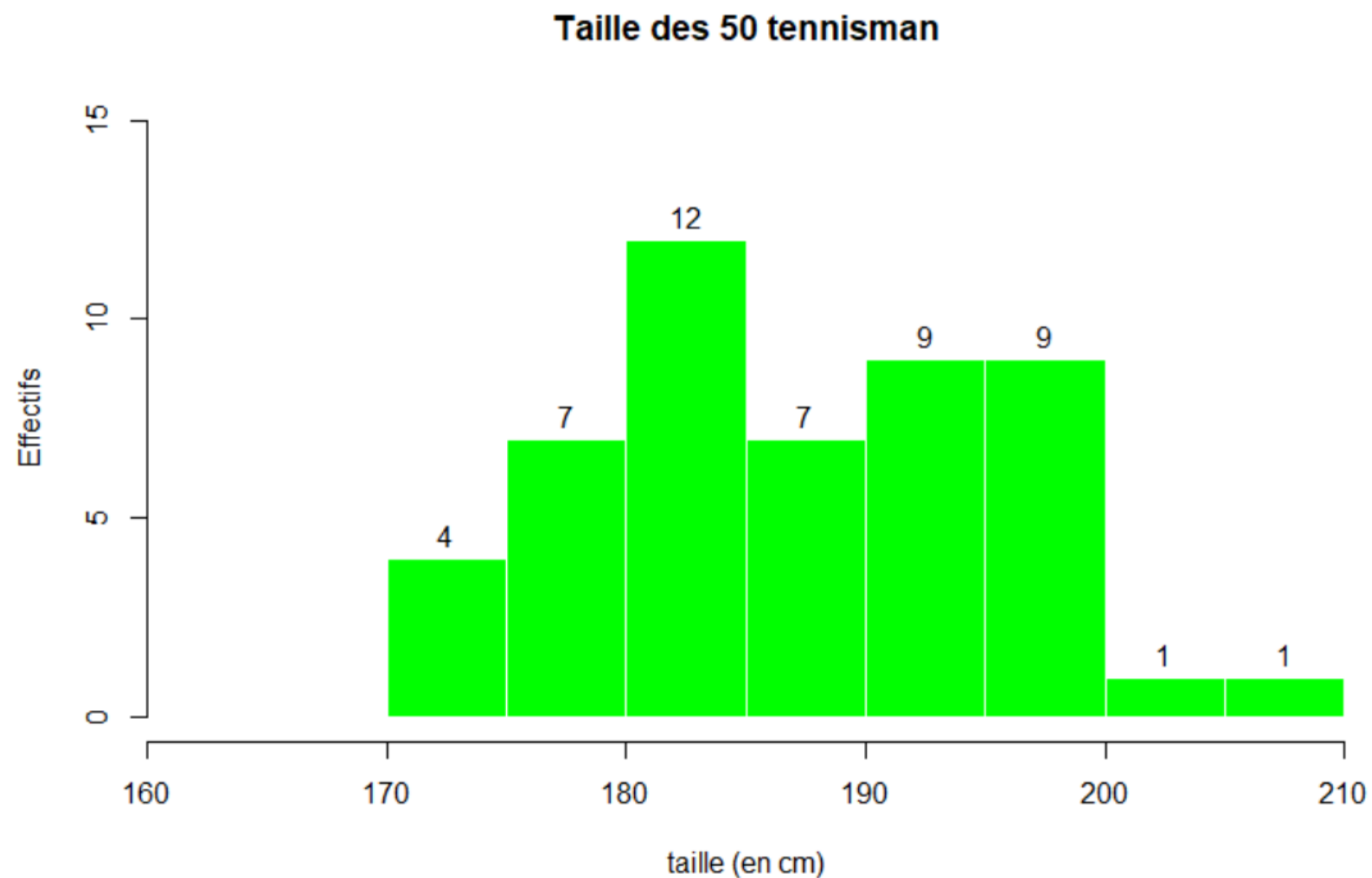
Ages des 50 tennisman



Moins de 26 ans	12
Entre 26 et 28 ans	9
Entre 29 et 31	19
Plus de 31	10

Min	20
1er quantile	26
Medianne	29
Moyenne	28.36
3eme quantile	31
Max	36

B)

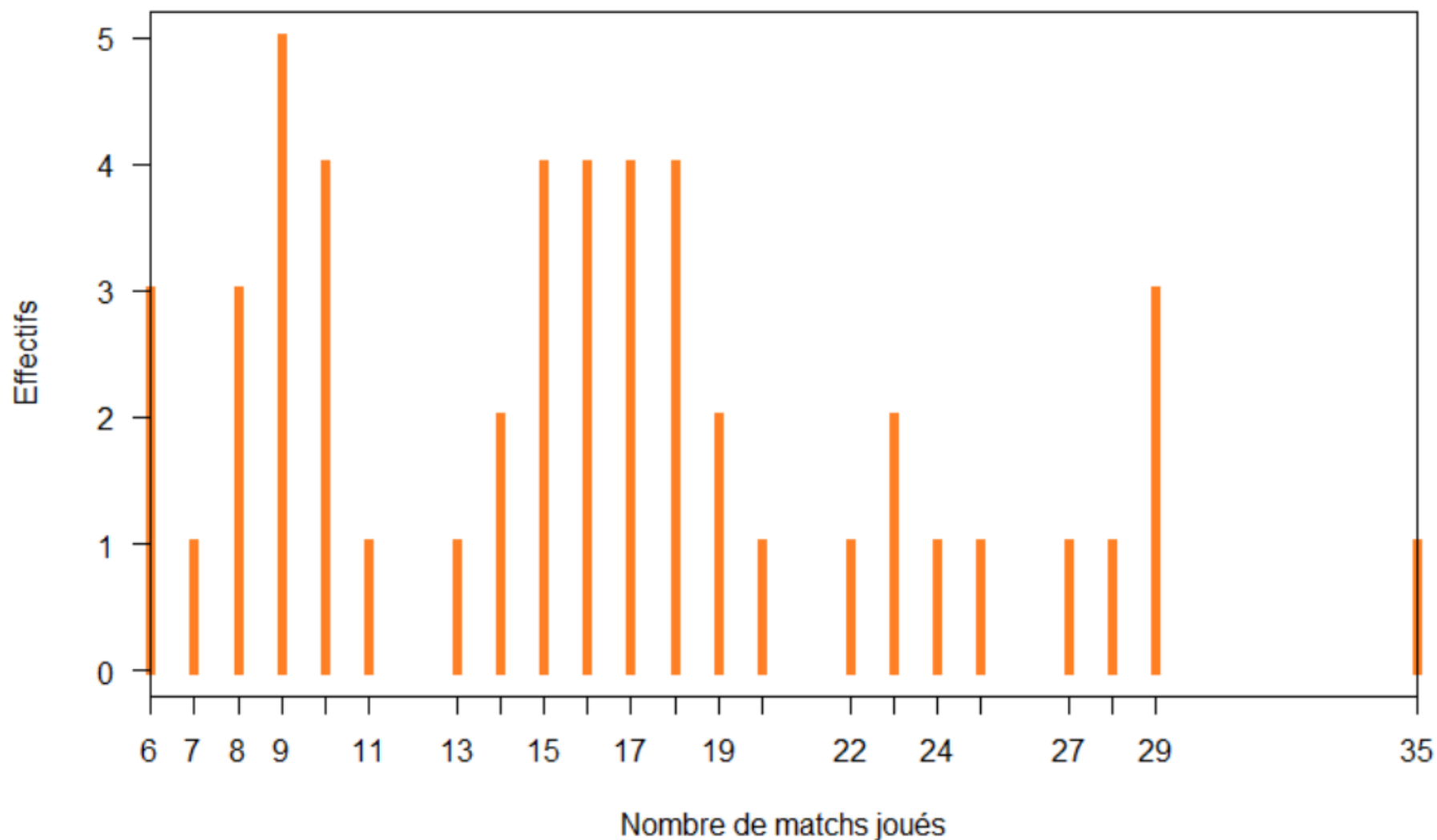


Moins de 180 cm	11
[180,190[21
[190,200[18
Plus de 200 cm	2

Min	170
1er quantile	183
Medianne	188
Moyenne	187.7
3eme quantile	193
Max	208

C)

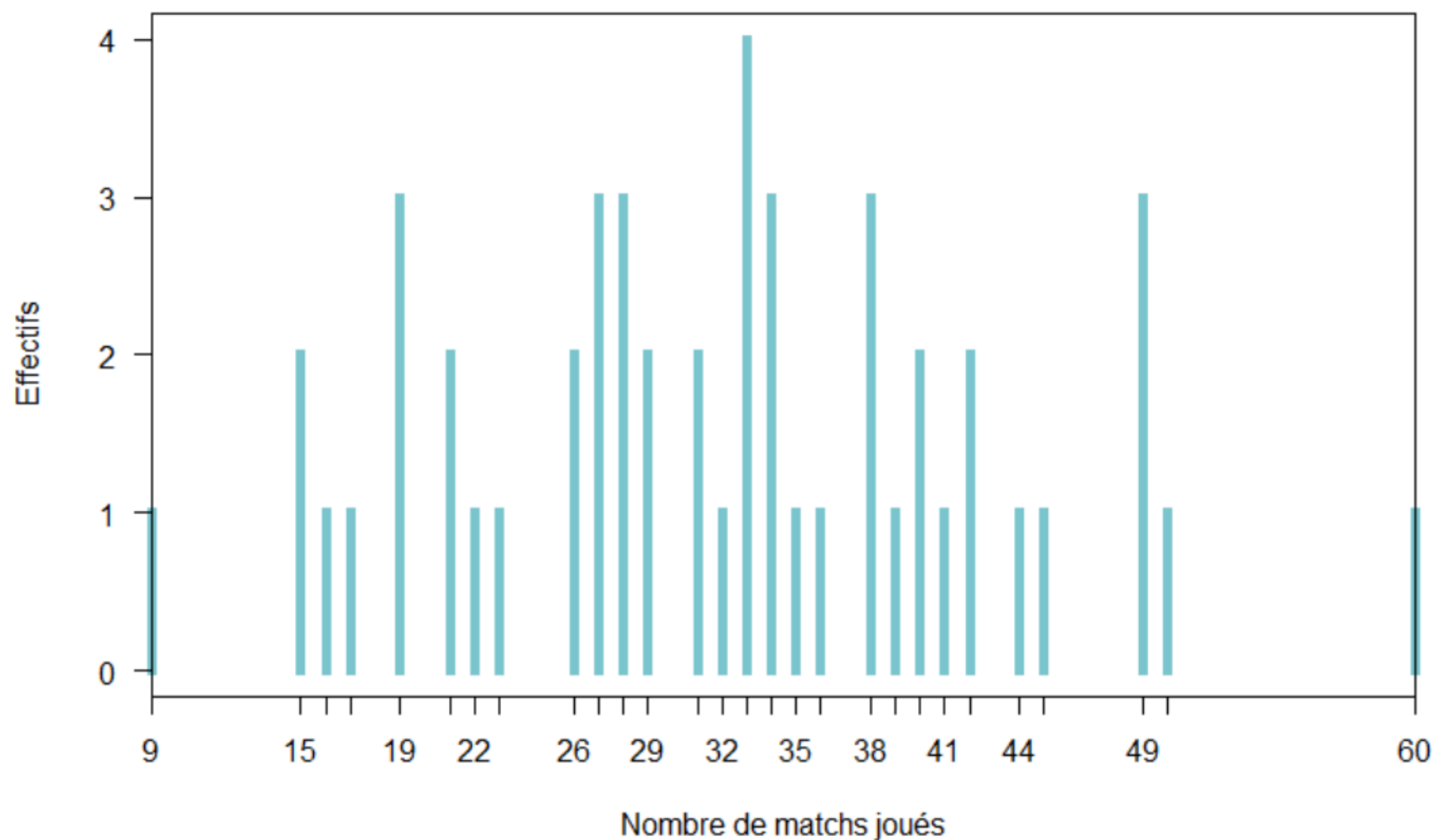
Nombre de matchs joués sur terre battue en 2017 par les 50 tennisman



Min	6
1er quantile	10
Medianne	16
Moyenne	16.04
3eme quantile	19
Max	35

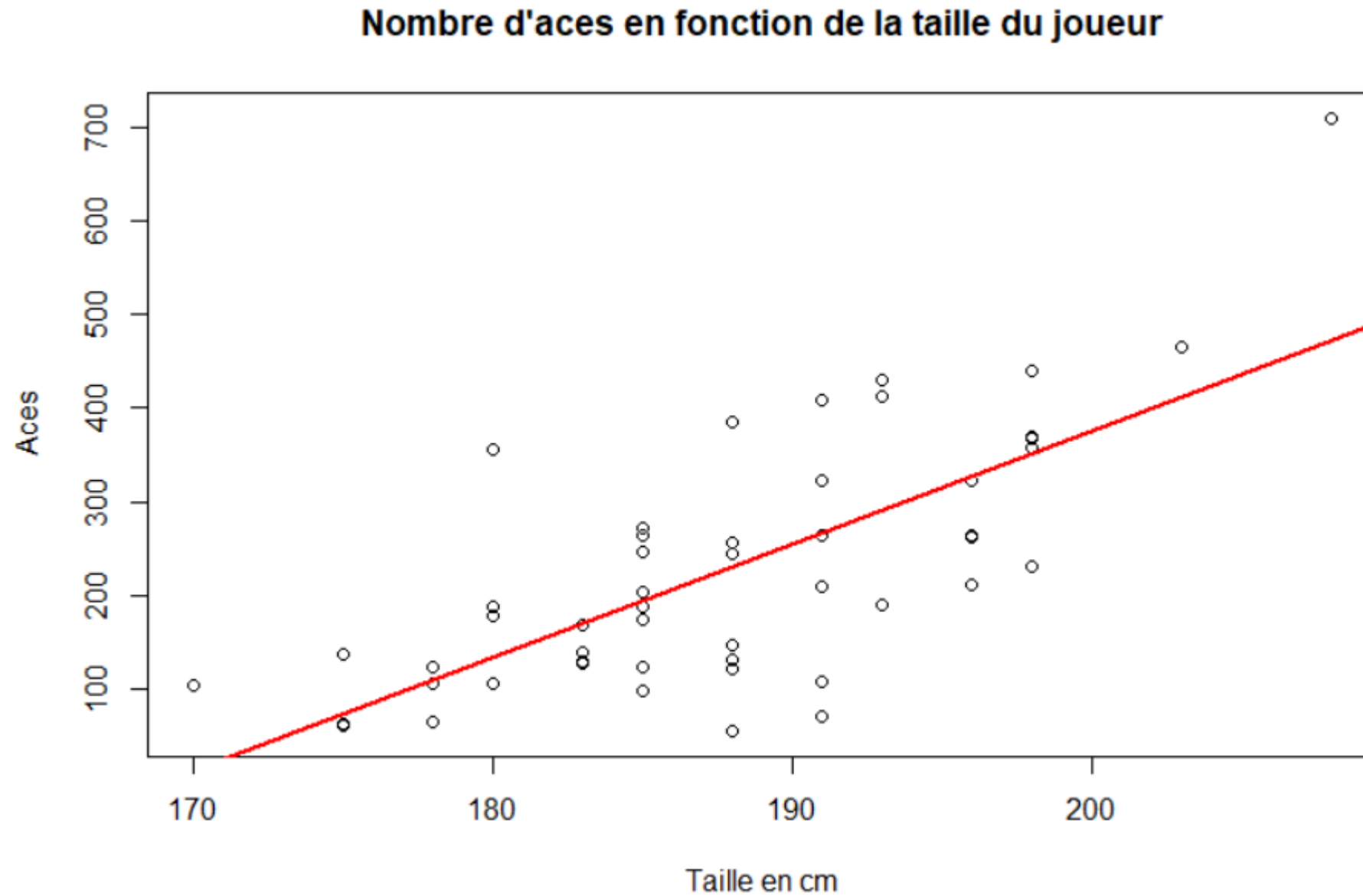
D)

Nombre de matchs joués sur dur en 2017 par les 50 tennisman



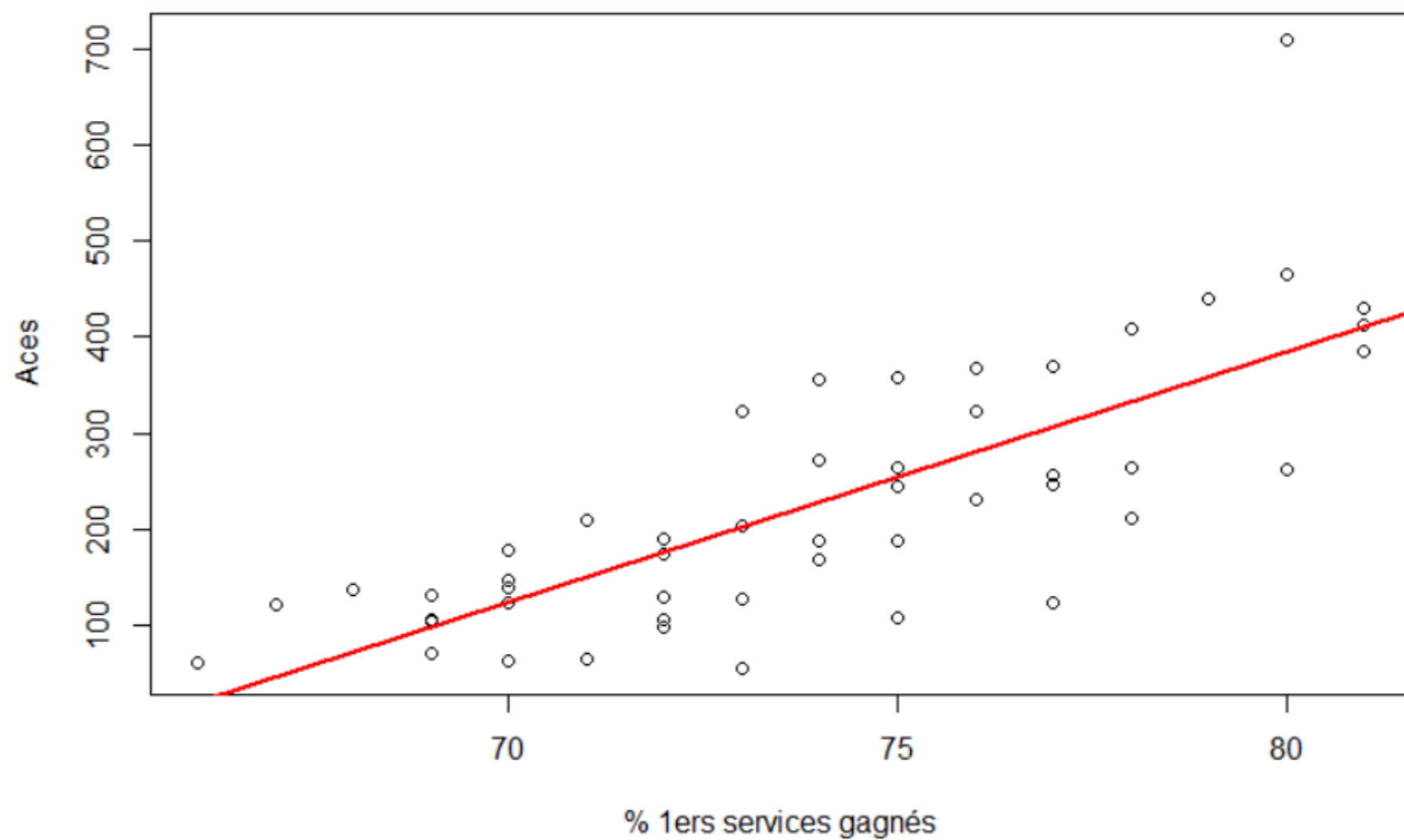
Min	9
1er quantile	26
Medianne	32.50
Moyenne	31.88
3eme quantile	38.75
Max	60

E)
On rappelle qu'un « ace » correspond à un service gagnant.



F)

Nombre d'aces en fonction du % de 1ers services gagnés



G)

X axis
Axis 1 (28.44%)

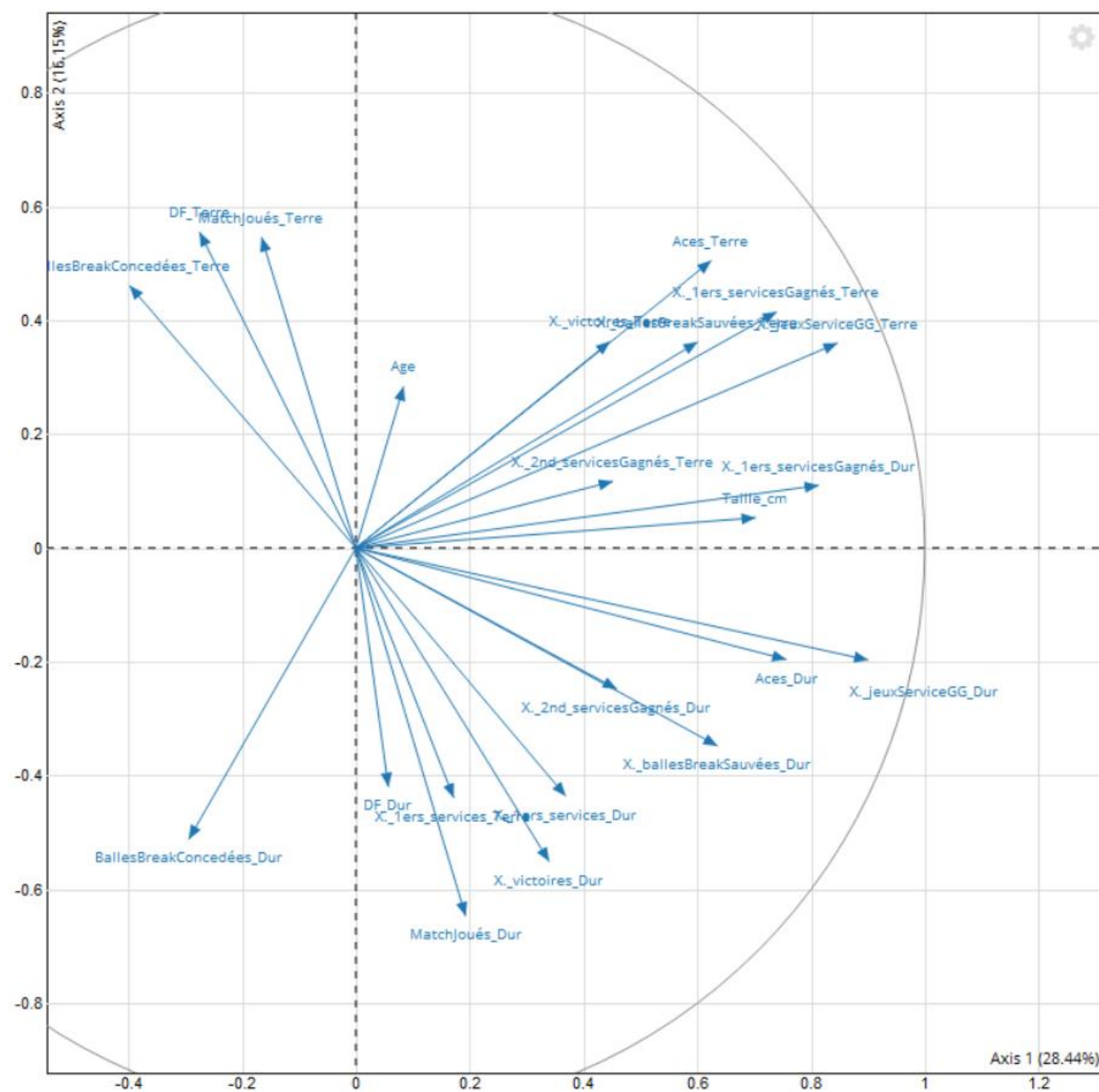
Y axis
Axis 2 (16.15%)

Labels size
4 10 20
4 6 8 10 12 16 20

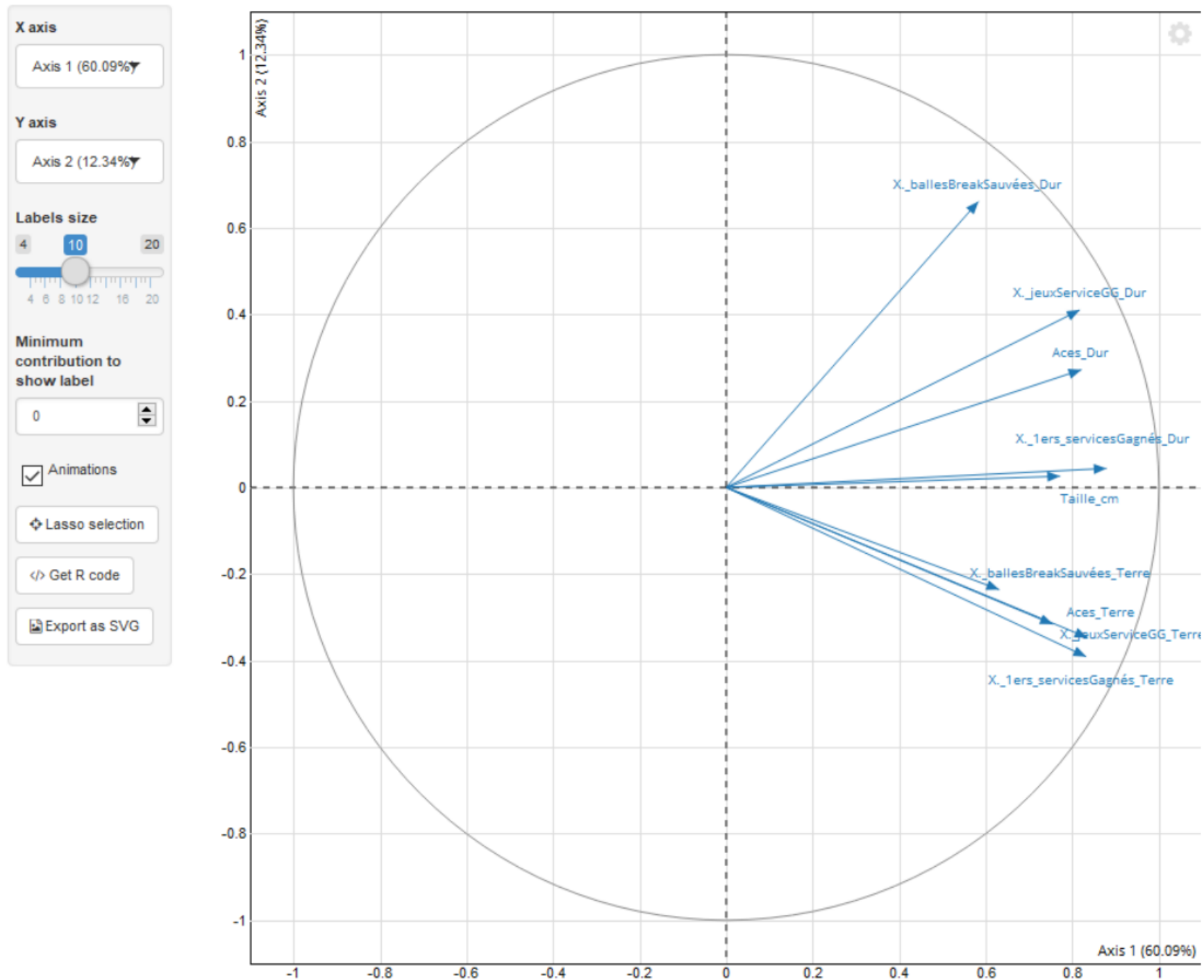
Minimum contribution to show label
0

☒ Animations

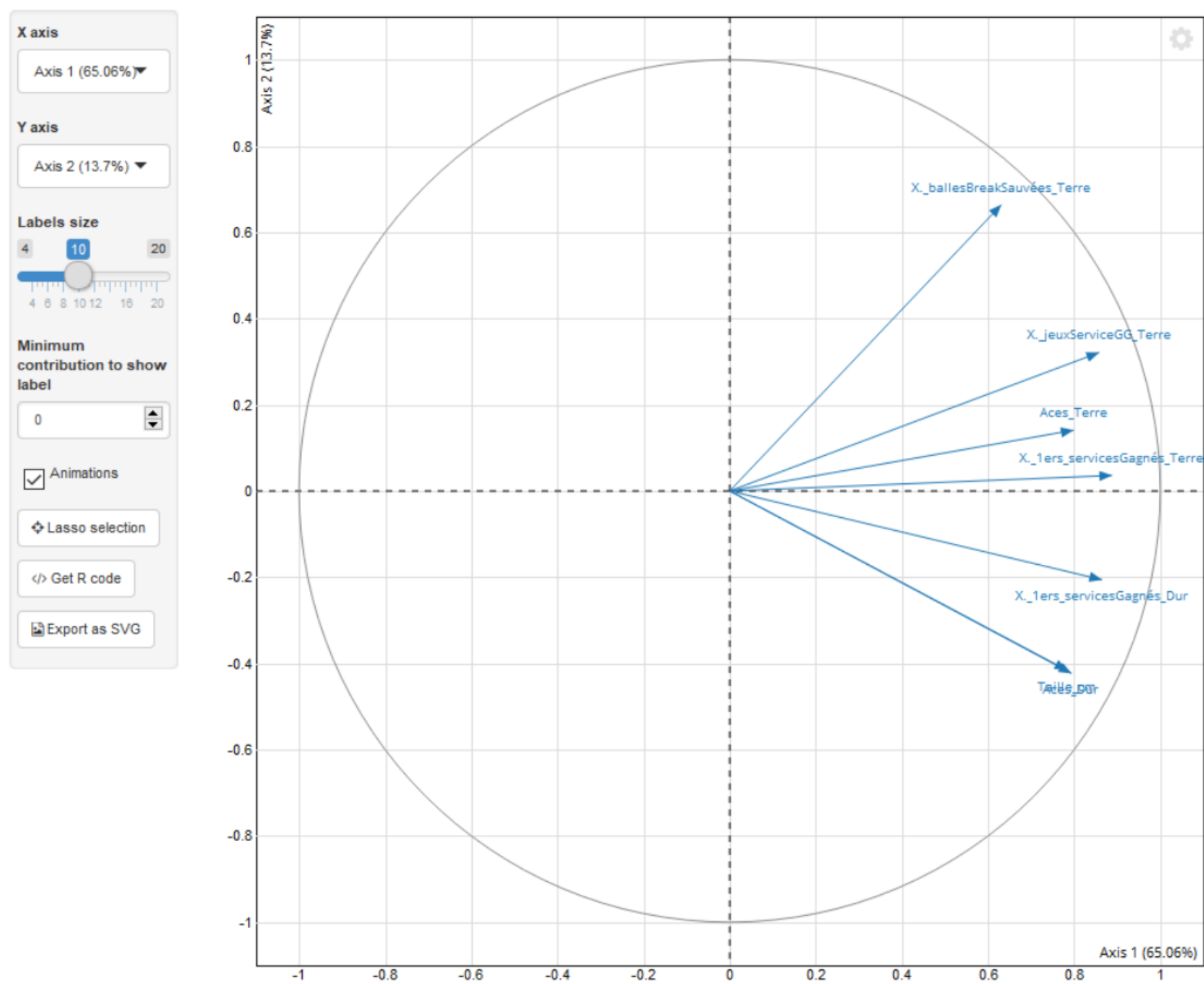
☒ Lasso selection



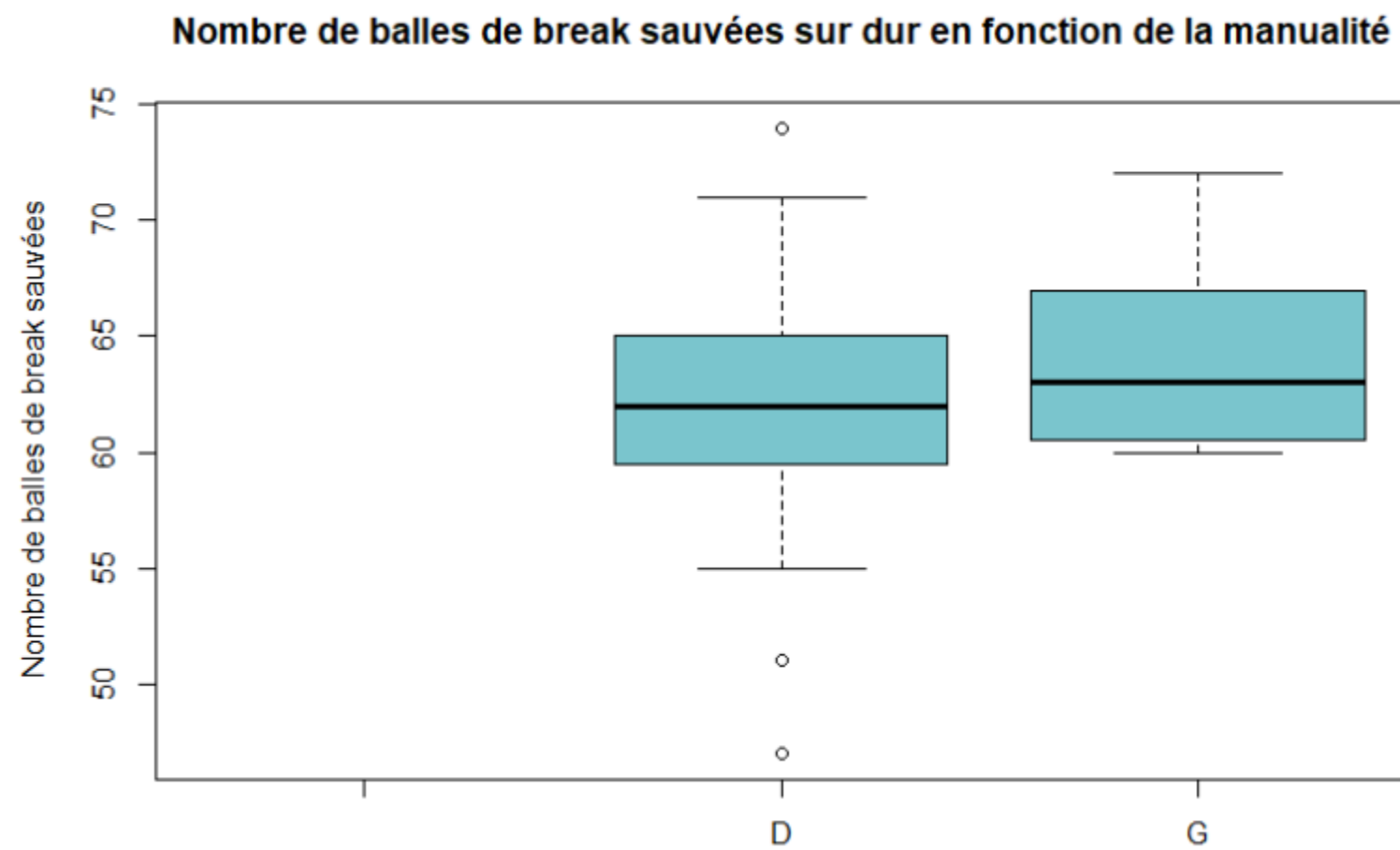
H)



1)

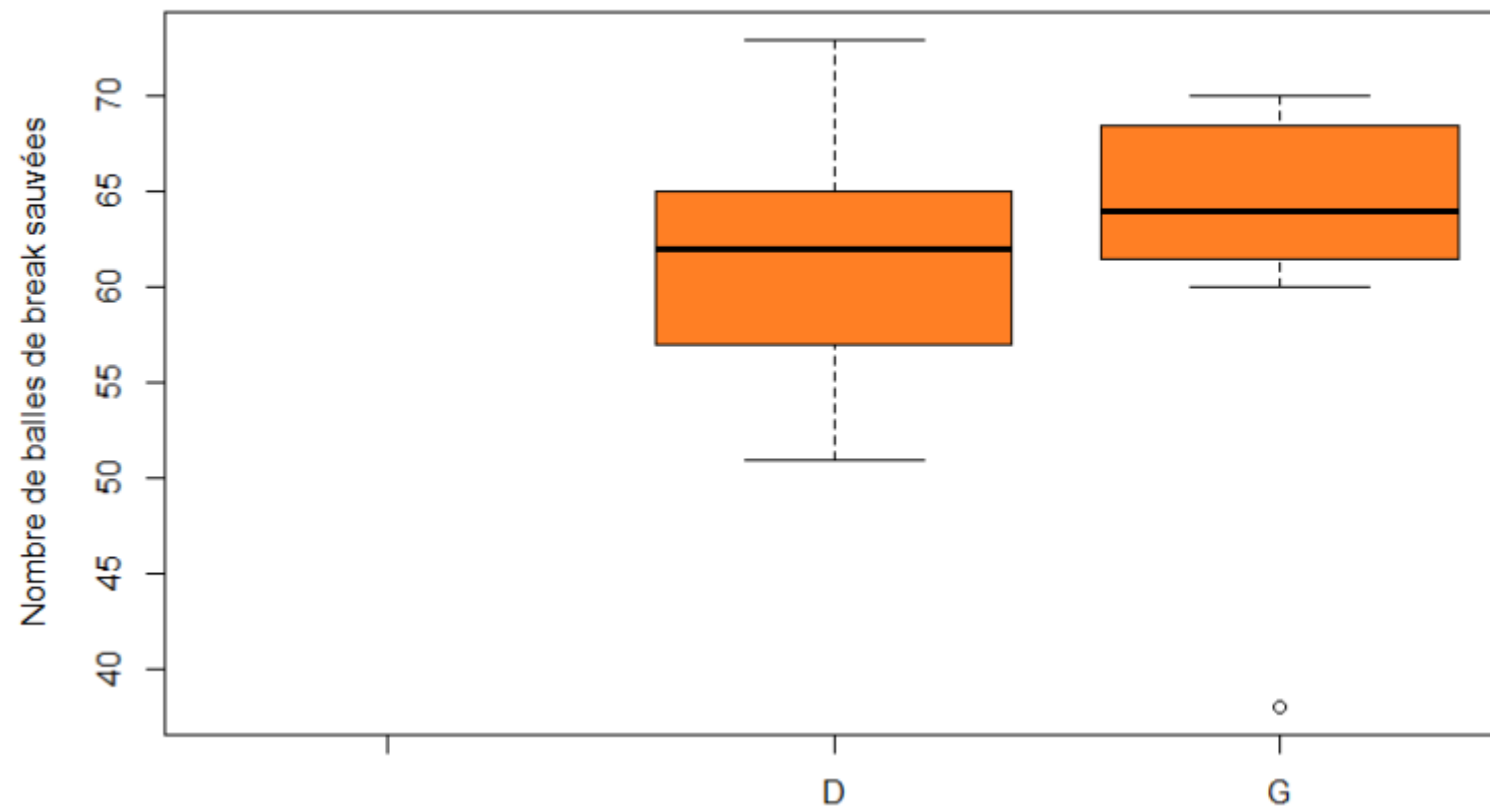


J)



K)

Nombre de balles de break sauvées sur terre battue en fonction de la manualité



Partie 2 : Synthèse d'une méthode vue en cours

Dans le cours, nous avons vu un ensemble de méthodes ayant pour but de croiser plusieurs variables.

Choisissez-une de ces méthodes et faites-en une synthèse comme si vous l'expliquiez à un élève ingénieur n'ayant pas eu de cours de Data Exploration.