

# Pràctica 2 - Neteja i anàlisi de les dades

David Navarro, Antoni Llussà

22/5/2020

## Contents

<b>Pràctica 2 - Neteja i anàlisi de les dades - Pokemon Dataset</b>	<b>2</b>
0. Carrega del dataset. . . . .	2
1. Descripció del dataset. Perquè és important i quina pregunta/problema pretén respondre? . . .	6
2. Integració i selecció de les dades d'interès a analitzar. . . . .	8
3. Neteja de les dades. . . . .	29
3.1. Les dades contenen zeros o elements buits? Com gestionaries aquests casos? . . . . .	29
3.2. Identificació i tractament de valors extrems. . . . .	30
4. Anàlisi de les dades. . . . .	31
4.1. Selecció dels grups de dades que es volen analitzar/comparar (planificació dels anàlisis a aplicar). . . . .	31
4.2. Comprovació de la normalitat i homogeneïtat de la variància. . . . .	33
4.3. Aplicació de proves estadístiques per comparar els grups de dades. En funció de les dades i de l'objectiu de l'estudi, aplicar proves de contrast d'hipòtesis, correlacions, regressions, etc. Aplicar almenys tres mètodes d'anàlisi diferents. . . . .	44
5. Representació dels resultats a partir de taules i gràfiques. . . . .	65
6. Resolució del problema. A partir dels resultats obtinguts, quines són les conclusions? Els resultats permeten respondre al problema? . . . . .	73
7. Codi: Cal adjuntar el codi, preferiblement en R, amb el que s'ha realitzat la neteja, anàlisi i representació de les dades. Si ho preferiu, també podeu treballar en Python. . . . .	74
8. Bibliografia . . . . .	74
9. Contribucions . . . . .	74

```
#carrega de llibreries per mostrar taules
library(knitr)
library(kableExtra)
library("ggplot2")
library(nortest)
library(corrplot)
```

```
## corrplot 0.84 loaded
```

## Pràctica 2 - Neteja i anàlisi de les dades - Pokemon Dataset

### 0. Carrega del dataset.

```
#pokedex <- read.csv("/Users/davidnavarrobrugal/Downloads/pokedex_(Update_05.20).csv",header=TRUE, stri
pokedex <- read.csv("dataset/pokedex_(Update_05.20).csv",header=TRUE, stringsAsFactors = TRUE,encoding :
```

Visualitzem les 5 primeres línies del csv.

```
head(pokedex)
```

```
##      X pokedex_number      name german_name
## 1 0          1      Bulbasaur      Bisasam
## 2 1          2      Ivysaur      Bisaknosp
## 3 2          3      Venusaur      Bisafloor
## 4 3          3 Mega Venusaur      Bisafloor
## 5 4          4      Charmander      Glumanda
## 6 5          5      Charmeleon      Glutexo
##
##                                     japanese_name generation status
## 1 <U+30D5><U+30B7><U+30AE><U+30C0><U+30CD> (Fushigidane)          1 Normal
## 2 <U+30D5><U+30B7><U+30AE><U+30BD><U+30A6> (Fushigisou)          1 Normal
## 3 <U+30D5><U+30B7><U+30AE><U+30D0><U+30CA> (Fushigibana)          1 Normal
## 4 <U+30D5><U+30B7><U+30AE><U+30D0><U+30CA> (Fushigibana)          1 Normal
## 5      <U+30D2><U+30C8><U+30AB><U+30B2> (Hitokage)              1 Normal
## 6      <U+30EA><U+30B6><U+30FC><U+30C9> (Lizardo)                1 Normal
##
##      species type_number type_1 type_2 height_m weight_kg abilities_number
## 1  Seed Pokémon          2  Grass Poison      0.7      6.9              2
## 2  Seed Pokémon          2  Grass Poison      1.0     13.0              2
## 3  Seed Pokémon          2  Grass Poison      2.0    100.0              2
## 4  Seed Pokémon          2  Grass Poison      2.4    155.5              1
## 5 Lizard Pokémon          1   Fire              0.6      8.5              2
## 6 Flame Pokémon          1   Fire              1.1     19.0              2
##
##      ability_1 ability_2 ability_hidden total_points hp attack defense sp_attack
## 1  Overgrow      Chlorophyll              318 45      49      49      65
## 2  Overgrow      Chlorophyll              405 60      62      63      80
## 3  Overgrow      Chlorophyll              525 80      82      83     100
## 4 Thick Fat              625 80     100     123     122
## 5  Blaze          Solar Power              309 39      52      43      60
## 6  Blaze          Solar Power              405 58      64      58      80
##
##      sp_defense speed catch_rate base_friendship base_experience growth_rate
## 1      65      45      45      70      64 Medium Slow
## 2      80      60      45      70     142 Medium Slow
## 3     100      80      45      70     236 Medium Slow
## 4     120      80      45      70     281 Medium Slow
## 5      50      65      45      70      62 Medium Slow
## 6      65      80      45      70     142 Medium Slow
##
##      egg_type_number egg_type_1 egg_type_2 percentage_male egg_cycles
## 1          2      Grass      Monster      87.5      20
## 2          2      Grass      Monster      87.5      20
## 3          2      Grass      Monster      87.5      20
## 4          2      Grass      Monster      87.5      20
## 5          2      Dragon      Monster      87.5      20
```

```
## 6          2      Dragon      Monster          87.5          20
##  against_normal against_fire against_water against_electric against_grass
## 1          1          2.0          0.5          0.5          0.25
## 2          1          2.0          0.5          0.5          0.25
## 3          1          2.0          0.5          0.5          0.25
## 4          1          1.0          0.5          0.5          0.25
## 5          1          0.5          2.0          1.0          0.50
## 6          1          0.5          2.0          1.0          0.50
##  against_ice against_fight against_poison against_ground against_flying
## 1          2.0          0.5          1          1          2
## 2          2.0          0.5          1          1          2
## 3          2.0          0.5          1          1          2
## 4          1.0          0.5          1          1          2
## 5          0.5          1.0          1          2          1
## 6          0.5          1.0          1          2          1
##  against_psychic against_bug against_rock against_ghost against_dragon
## 1          2          1.0          1          1          1
## 2          2          1.0          1          1          1
## 3          2          1.0          1          1          1
## 4          2          1.0          1          1          1
## 5          1          0.5          2          1          1
## 6          1          0.5          2          1          1
##  against_dark against_steel against_fairy
## 1          1          1.0          0.5
## 2          1          1.0          0.5
## 3          1          1.0          0.5
## 4          1          1.0          0.5
## 5          1          0.5          0.5
## 6          1          0.5          0.5
```

Per inspeccionar el fitxer, utilitzarem la instrucció *str*. On veurem el nombre d'observacions, el nombre de variables, el nom de les variables i de quin tipus són.

```
str(pokedex)
```

```
## 'data.frame':    1028 obs. of  51 variables:
## $ X              : int  0 1 2 3 4 5 6 7 8 9 ...
## $ pokedex_number : int  1 2 3 3 4 5 6 6 6 7 ...
## $ name           : Factor w/ 1028 levels "Abomasnow","Abra",...: 105 426 958 582 132 133 130 548 548 ...
## $ german_name    : Factor w/ 808 levels "", "Aalabyss",...: 54 53 52 52 221 223 222 222 222 607 ...
## $ japanese_name  : Factor w/ 808 levels "", "<U+30A2><U+30A4><U+30A2><U+30F3><U+30C8> (Aiant)",...: 105 426 958 582 132 133 130 548 548 ...
## $ generation     : int  1 1 1 1 1 1 1 1 1 1 ...
## $ status         : Factor w/ 4 levels "Legendary","Mythical",...: 3 3 3 3 3 3 3 3 3 3 ...
## $ species        : Factor w/ 641 levels "Abundance Pokémon",...: 487 487 487 487 328 208 208 208 208 208 ...
## $ type_number    : int  2 2 2 2 1 1 2 2 2 1 ...
## $ type_1         : Factor w/ 18 levels "Bug","Dark","Dragon",...: 10 10 10 10 7 7 7 7 7 18 ...
## $ type_2         : Factor w/ 19 levels "", "Bug","Dark",...: 15 15 15 15 1 1 9 4 9 1 ...
## $ height_m       : num  0.7 1 2 2.4 0.6 1.1 1.7 1.7 1.7 0.5 ...
## $ weight_kg      : num  6.9 13 100 155.5 8.5 ...
## $ abilities_number: int  2 2 2 1 2 2 2 1 1 2 ...
## $ ability_1      : Factor w/ 203 levels "", "Adaptability",...: 110 110 110 184 16 16 16 186 41 185 ...
## $ ability_2      : Factor w/ 127 levels "", "Adaptability",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ ability_hidden  : Factor w/ 155 levels "", "Adaptability",...: 12 12 12 1 121 121 121 1 1 98 ...
```

```
## $ total_points : num 318 405 525 625 309 405 534 634 634 314 ...
## $ hp : num 45 60 80 80 39 58 78 78 78 44 ...
## $ attack : num 49 62 82 100 52 64 84 130 104 48 ...
## $ defense : num 49 63 83 123 43 58 78 111 78 65 ...
## $ sp_attack : num 65 80 100 122 60 80 109 130 159 50 ...
## $ sp_defense : num 65 80 100 120 50 65 85 85 115 64 ...
## $ speed : num 45 60 80 80 65 80 100 100 100 43 ...
## $ catch_rate : num 45 45 45 45 45 45 45 45 45 45 ...
## $ base_friendship : num 70 70 70 70 70 70 70 70 70 70 ...
## $ base_experience : num 64 142 236 281 62 142 240 285 285 63 ...
## $ growth_rate : Factor w/ 7 levels "", "Erratic", "Fast", ...: 6 6 6 6 6 6 6 6 6 6 ...
## $ egg_type_number : int 2 2 2 2 2 2 2 2 2 2 ...
## $ egg_type_1 : Factor w/ 16 levels "", "Amorphous", ...: 9 9 9 9 5 5 5 5 12 ...
## $ egg_type_2 : Factor w/ 12 levels "", "Dragon", "Fairy", ...: 9 9 9 9 9 9 9 9 10 ...
## $ percentage_male : num 87.5 87.5 87.5 87.5 87.5 87.5 87.5 87.5 87.5 87.5 ...
## $ egg_cycles : num 20 20 20 20 20 20 20 20 20 20 ...
## $ against_normal : num 1 1 1 1 1 1 1 1 1 1 ...
## $ against_fire : num 2 2 2 1 0.5 0.5 0.5 0.25 0.5 0.5 ...
## $ against_water : num 0.5 0.5 0.5 0.5 2 2 2 1 2 0.5 ...
## $ against_electric : num 0.5 0.5 0.5 0.5 1 1 2 0.5 2 2 ...
## $ against_grass : num 0.25 0.25 0.25 0.25 0.5 0.5 0.25 0.25 0.25 2 ...
## $ against_ice : num 2 2 2 1 0.5 0.5 1 1 1 0.5 ...
## $ against_fight : num 0.5 0.5 0.5 0.5 1 1 0.5 1 0.5 1 ...
## $ against_poison : num 1 1 1 1 1 1 1 1 1 1 ...
## $ against_ground : num 1 1 1 1 2 2 0 2 0 1 ...
## $ against_flying : num 2 2 2 2 1 1 1 1 1 1 ...
## $ against_psychic : num 2 2 2 2 1 1 1 1 1 1 ...
## $ against_bug : num 1 1 1 1 0.5 0.5 0.25 0.5 0.25 1 ...
## $ against_rock : num 1 1 1 1 2 2 4 2 4 1 ...
## $ against_ghost : num 1 1 1 1 1 1 1 1 1 1 ...
## $ against_dragon : num 1 1 1 1 1 1 1 2 1 1 ...
## $ against_dark : num 1 1 1 1 1 1 1 1 1 1 ...
## $ against_steel : num 1 1 1 1 0.5 0.5 0.5 0.5 0.5 0.5 ...
## $ against_fairy : num 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 1 0.5 1 ...
```

Utilitzarem la instrucció *summary* per veure les freqüències, min, max, mitjan, mediana, quartils i valors absents.

```
summary(pokedex)
```

```
##           X          pokedex_number          name
## Min.      : 0.0    Min.      : 1.0    Abomasnow      : 1
## 1st Qu.: 256.8    1st Qu.:213.8    Abra            : 1
## Median : 513.5    Median :433.5    Absol          : 1
## Mean    : 513.5    Mean    :437.7    Accelgor       : 1
## 3rd Qu.: 770.2    3rd Qu.:663.2    Aegislash Blade Forme : 1
## Max.    :1027.0    Max.    :890.0    Aegislash Shield Forme: 1
##                                     (Other)          :1022
##      german_name                                     japanese_name
##           : 90                                     : 90
## Rotom      : 6    <U+30ED><U+30C8><U+30E0> (Rotom)      : 6
## Choreogel  : 4    <U+30AA><U+30C9><U+30EA><U+30C9><U+30EA> (Odoridori): 4
## Deoxys     : 4    <U+30C7><U+30AA><U+30AD><U+30B7><U+30B9> (Deoxys)   : 4
```

```

## Flampivian: 4 <U+30CD><U+30AF><U+30ED><U+30BA><U+30DE> (Necrozma) : 4
## Formeo : 4 <U+30D0><U+30B1><U+30C3><U+30C1><U+30E3> (Bakeccha) : 4
## (Other) :916 (Other) :916
## generation status species type_number
## Min. :1.000 Legendary : 39 Mouse Pokémon : 12 Min. :1.000
## 1st Qu.:2.000 Mythical : 29 Dragon Pokémon : 9 1st Qu.:1.000
## Median :4.000 Normal :915 Fox Pokémon : 9 Median :2.000
## Mean :4.034 Sub Legendary: 45 Pumpkin Pokémon: 8 Mean :1.527
## 3rd Qu.:6.000 Flame Pokémon : 7 3rd Qu.:2.000
## Max. :8.000 Bagworm Pokémon: 6 Max. :2.000
## (Other) :977
## type_1 type_2 height_m weight_kg
## Water :134 :486 Min. : 0.100 Min. : 0.10
## Normal :115 Flying :109 1st Qu.: 0.600 1st Qu.: 8.80
## Grass : 91 Fairy : 41 Median : 1.000 Median : 28.50
## Bug : 81 Ground : 39 Mean : 1.368 Mean : 69.75
## Psychic: 76 Poison : 38 3rd Qu.: 1.500 3rd Qu.: 69.10
## Fire : 65 Psychic: 38 Max. :100.000 Max. :999.90
## (Other):466 (Other):277 NA's :1
## abilities_number ability_1 ability_2 ability_hidden
## Min. :0.000 Levitate : 41 :515 :218
## 1st Qu.:2.000 Swift Swim : 30 Frisk : 17 Telepathy : 21
## Median :2.000 Keen Eye : 28 Sturdy : 14 Sheer Force: 18
## Mean :2.284 Chlorophyll: 25 Shell Armor: 13 Overcoat : 17
## 3rd Qu.:3.000 Pressure : 25 Early Bird : 10 Unnerve : 17
## Max. :3.000 Blaze : 24 Gluttony : 10 Regenerator: 15
## (Other) :855 (Other) :449 (Other) :722
## total_points hp attack defense
## Min. : 175.0 Min. : 1.00 Min. : 5.00 Min. : 5.00
## 1st Qu.: 330.0 1st Qu.: 50.00 1st Qu.: 55.00 1st Qu.: 50.00
## Median : 455.0 Median : 66.50 Median : 76.00 Median : 70.00
## Mean : 437.6 Mean : 69.58 Mean : 80.12 Mean : 74.48
## 3rd Qu.: 510.0 3rd Qu.: 80.00 3rd Qu.:100.00 3rd Qu.: 90.00
## Max. :1125.0 Max. :255.00 Max. :190.00 Max. :250.00
##
## sp_attack sp_defense speed catch_rate
## Min. : 10.00 Min. : 20.00 Min. : 5.00 Min. : 3.00
## 1st Qu.: 50.00 1st Qu.: 50.00 1st Qu.: 45.00 1st Qu.: 45.00
## Median : 65.00 Median : 70.00 Median : 65.00 Median : 60.00
## Mean : 72.73 Mean : 72.13 Mean : 68.53 Mean : 93.17
## 3rd Qu.: 95.00 3rd Qu.: 90.00 3rd Qu.: 90.00 3rd Qu.:127.00
## Max. :194.00 Max. :250.00 Max. :180.00 Max. :255.00
## NA's :104
## base_friendship base_experience growth_rate egg_type_number
## Min. : 0.00 Min. : 36.0 : 1 Min. :0.000
## 1st Qu.: 70.00 1st Qu.: 67.0 Erratic : 26 1st Qu.:1.000
## Median : 70.00 Median :159.0 Fast : 68 Median :1.000
## Mean : 64.14 Mean :153.8 Fluctuating: 14 Mean :1.271
## 3rd Qu.: 70.00 3rd Qu.:201.5 Medium Fast:432 3rd Qu.:2.000
## Max. :140.00 Max. :608.0 Medium Slow:245 Max. :2.000
## NA's :104 NA's :104 Slow :242
## egg_type_1 egg_type_2 percentage_male egg_cycles
## Field :258 :746 Min. : 0 Min. : 5.00
## Undiscovered:138 Water 1 : 76 1st Qu.: 50 1st Qu.: 20.00

```

```

## Bug      : 90   Monster : 55   Median : 50   Median : 20.00
## Amorphous : 84   Human-Like: 28   Mean   : 55   Mean   : 30.32
## Dragon    : 67   Grass     : 27   3rd Qu.: 50   3rd Qu.: 25.00
## Fairy     : 62   Field     : 21   Max.    :100   Max.    :120.00
## (Other)   :329   (Other)   : 75   NA's    :236   NA's    :1
## against_normal  against_fire  against_water  against_electric
## Min.    :0.0000  Min.    :0.000  Min.    :0.000  Min.    :0.000
## 1st Qu.:1.0000  1st Qu.:0.500  1st Qu.:0.500  1st Qu.:0.500
## Median :1.0000  Median :1.000  Median :1.000  Median :1.000
## Mean    :0.8684  Mean    :1.125  Mean    :1.054  Mean    :1.034
## 3rd Qu.:1.0000  3rd Qu.:2.000  3rd Qu.:1.000  3rd Qu.:1.000
## Max.    :1.0000  Max.    :4.000  Max.    :4.000  Max.    :4.000
##
## against_grass   against_ice   against_fight  against_poison
## Min.    :0.000  Min.    :0.000  Min.    :0.000  Min.    :0.0000
## 1st Qu.:0.500  1st Qu.:0.500  1st Qu.:0.500  1st Qu.:0.5000
## Median :1.000  Median :1.000  Median :1.000  Median :1.0000
## Mean    :1.004  Mean    :1.196  Mean    :1.079  Mean    :0.9523
## 3rd Qu.:1.000  3rd Qu.:2.000  3rd Qu.:2.000  3rd Qu.:1.0000
## Max.    :4.000  Max.    :4.000  Max.    :4.000  Max.    :4.0000
##
## against_ground  against_flying  against_psychic  against_bug
## Min.    :0.000  Min.    :0.250  Min.    :0.0000  Min.    :0.0000
## 1st Qu.:0.500  1st Qu.:1.000  1st Qu.:1.0000  1st Qu.:0.5000
## Median :1.000  Median :1.000  Median :1.0000  Median :1.0000
## Mean    :1.085  Mean    :1.166  Mean    :0.9793  Mean    :0.9925
## 3rd Qu.:1.625  3rd Qu.:1.000  3rd Qu.:1.0000  3rd Qu.:1.0000
## Max.    :4.000  Max.    :4.000  Max.    :4.0000  Max.    :4.0000
##
## against_rock    against_ghost  against_dragon  against_dark
## Min.    :0.25  Min.    :0.000  Min.    :0.0000  Min.    :0.250
## 1st Qu.:1.00  1st Qu.:1.000  1st Qu.:1.0000  1st Qu.:1.000
## Median :1.00  Median :1.000  Median :1.0000  Median :1.000
## Mean    :1.24  Mean    :1.011  Mean    :0.9757  Mean    :1.066
## 3rd Qu.:2.00  3rd Qu.:1.000  3rd Qu.:1.0000  3rd Qu.:1.000
## Max.    :4.00  Max.    :4.000  Max.    :2.0000  Max.    :4.000
##
## against_steel    against_fairy
## Min.    :0.0000  Min.    :0.000
## 1st Qu.:0.5000  1st Qu.:1.000
## Median :1.0000  Median :1.000
## Mean    :0.9803  Mean    :1.085
## 3rd Qu.:1.0000  3rd Qu.:1.000
## Max.    :4.0000  Max.    :4.000
##

```

## 1. Descripció del dataset. Perquè és important i quina pregunta/problema pretén respondre?

El dataset escollit, és una base de dades del joc de Pokemon. El dataset, es pot descarregar a: <https://www.kaggle.com/mariotormo/complete-pokemon-dataset-updated-090420>, té la Llicència CC BY-SA 4.0, per tant la podem usar per realitzar en aquesta pràctica.

S'ha escollit aquest dataset perquè els dos membres del grup, tenim fills, i estan aficionats en aquest món.

Aquest dataset ens ajudarà a saber quines tipologies de Pokemon hi ha, quines característiques tenen, quins són els Pokemons més forts, etc. La nostra experiència amb el joc, és nul·la, per tant haurem d'analitzar profundament el dataset i entendre que és cada una de les variables.

El dataset escollit, té dos arxius .csv, aquests dos arxius, estan actualitzats en dates diferents. Els hem analitzat i tots contenen la mateixa informació, però estructurada diferent per mostrar si són: Normal, Legendary, Mythical, Sub Legendary.

Utilitzarem l'arxiu `pokedex_(Update_05.20).csv`, conté 1028 observacions i 51 variables. La variable X, no la utilitzarem ja que no ens aporta informació, és un identificador de fila.

Les dades es poden agrupar en:

#### 1. Informació de la Pokedex:

- *Pokedex\_number*: Identificador del Pokémon dins de la Pokedex nacional
- *name*: Nom del Pokémon amb anglès.
- *german\_name*: Nom del Pokémon amb alemany.
- *japanese\_name*: Nom del Pokémon amb japonès.
- *generation*: Número de generació en que s'ha vist per primera vegada.
- *status*: Indica si el Pokémon és: normal, sub legendary, legendary or mythical
- *species*: Categoria del Pokémon.
- *type\_number*: Nombre de tipus que té el Pokémon.
- *type\_1*: El tipus primari del Pokémon.
- *type\_2*: El tipus secundari del Pokémon.
- *height\_m*: Alçada del Pokémon en metres.
- *weight\_kg*: Pes del Pokémon en kg.
- *abilities\_number*: Nombre d'habilitats del Pokémon.
- *ability\_?*: Nom de l'habilitat.
- *ability\_hidden*: Si té habilitat oculta, Nom de l'habilitat del Pokémon.

#### 2. Informació d'estadístiques bàsiques:

- *total\_points*: Nombre total de punts de base.
- *hp*: L'HP de base del Pokémon.
- *attack*: L'atac de base del Pokémon.
- *defense*: La defensa base del Pokémon.
- *sp\_attack*: Atac especial de base del Pokémon.
- *sp\_defense*: Defensa especial de base del Pokémon.
- *speed*: La velocitat base del Pokémon.

#### 3. Informació d'entrenament:

- *catch\_rate*: Taxa de captura del Pokémon.
- *base\_friendship*: L'amistat base del Pokémon.
- *base\_experience*: L'experiència de base d'un Pokémon salvatge quan és capturat.
- *growth\_rate*: La taxa de creixement del Pokémon.

#### 4. Informació de cria:

- *egg\_type\_number*: Nombre de grups on pot ser eclosionat.
- *egg\_type\_?*: Noms dels grups d'ou on es pot eclosionar un Pokemon.
- *percentage\_male*: El percentatge de les espècies que són mascles. En blanc si el Pokemon no té gènere.

- *egg\_cycles*: El nombre de cicles (255-257 passos) necessaris per eclosionar un ou del Pokémon.

5. Informació de tipus de defensa:

- *against\_?*: Divuit funcions que informen la quantitat de danys causats contra un atac d'un tipus particular.

## 2. Integració i selecció de les dades d'interès a analitzar.

En aquest cas el dataset escollit, no es pot realitzar cap integració amb altres datasets, no n'hem trobat cap que pogués aportar més valor.

Per saber les dades d'interès a analitzar realitzarem un anàlisi visual per veure les variables de tipus factor i numèriques que ens criden l'atenció, per poder llavors decidir quin es l'estudi que es realitzarà:

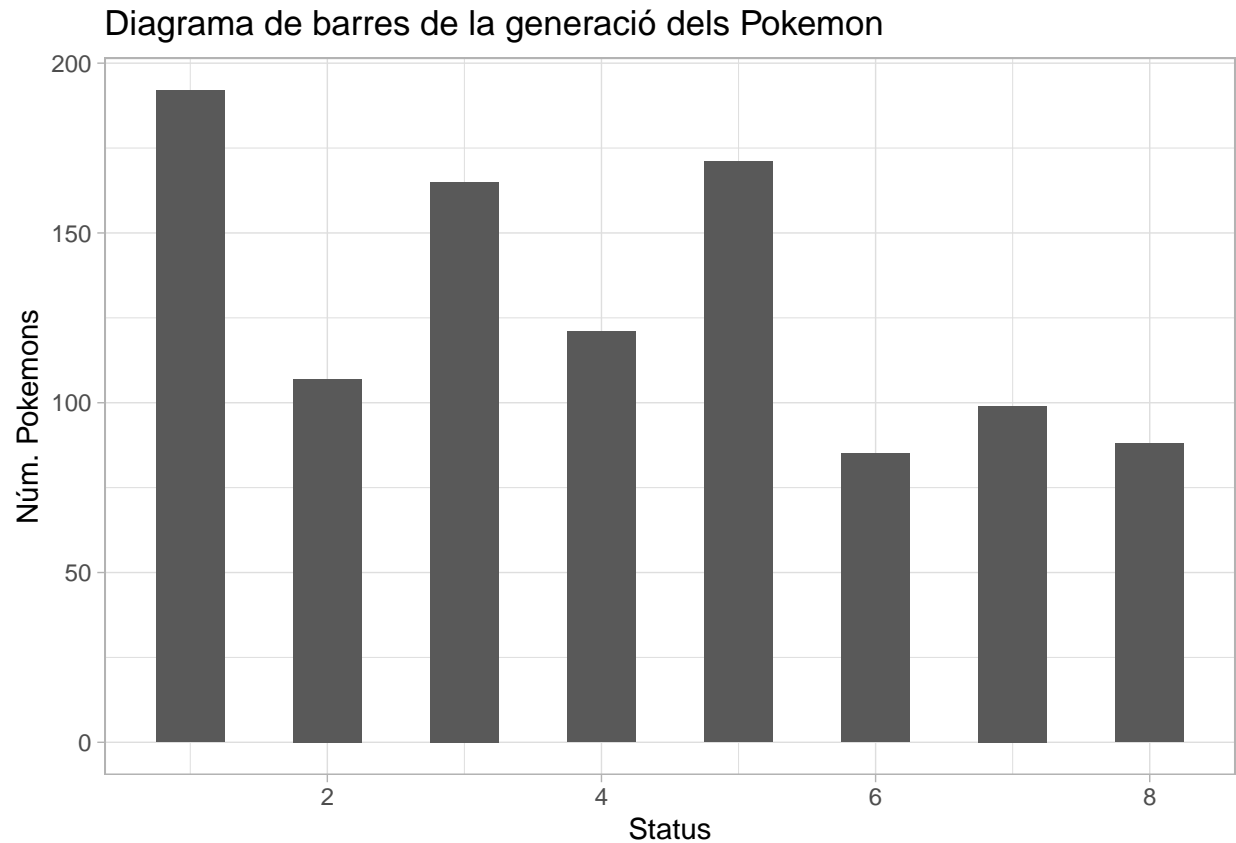
### Anàlisi descriptiva visual

#### Anàlisi de variables de tipus factor

```
gg <- ggplot(pokedex, aes(x = generation, fill = generation)) +
  geom_bar(width=0.5) + # bar chart
  theme_light() + # theme
  labs(x = "Status", # this labels the x axis
       y = "Núm. Pokemons", # this labels the y axis
       title = "Diagrama de barres de la generació dels Pokemon") # title of the plot
gg
```

#### Anàlisi de la variables generation

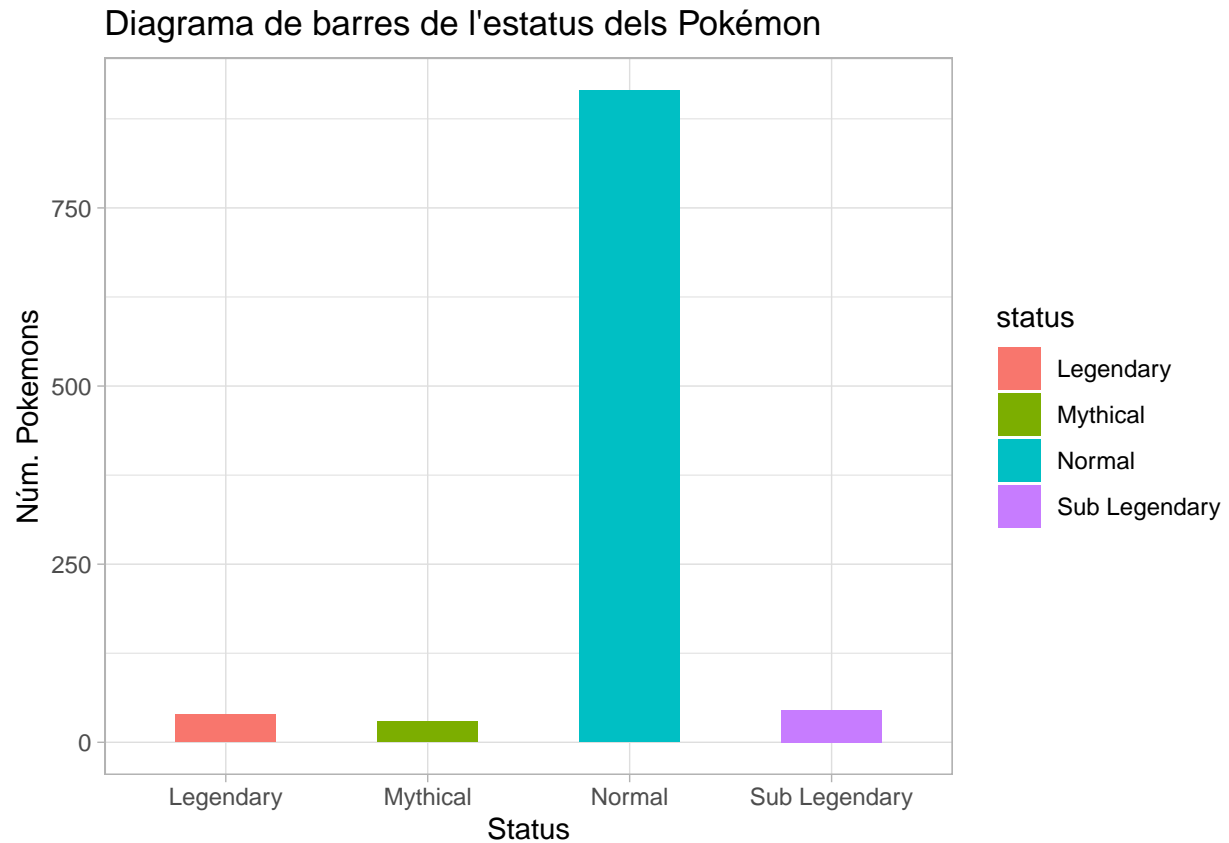




Es pot observar que a la primera generació és la que van aparèixer més Pokemons, llavors la cinquena i la tercera.

```
gg <- ggplot(pokedex, aes(x = status, fill = status)) +  
  geom_bar(width=0.5) + # bar chart  
  theme_light() + # theme  
  labs(x = "Status", # this labels the x axis  
       y = "Núm. Pokemons", # this labels the y axis  
       title = "Diagrama de barres de l'estatus dels Pokémon") # title of the plot  
gg
```

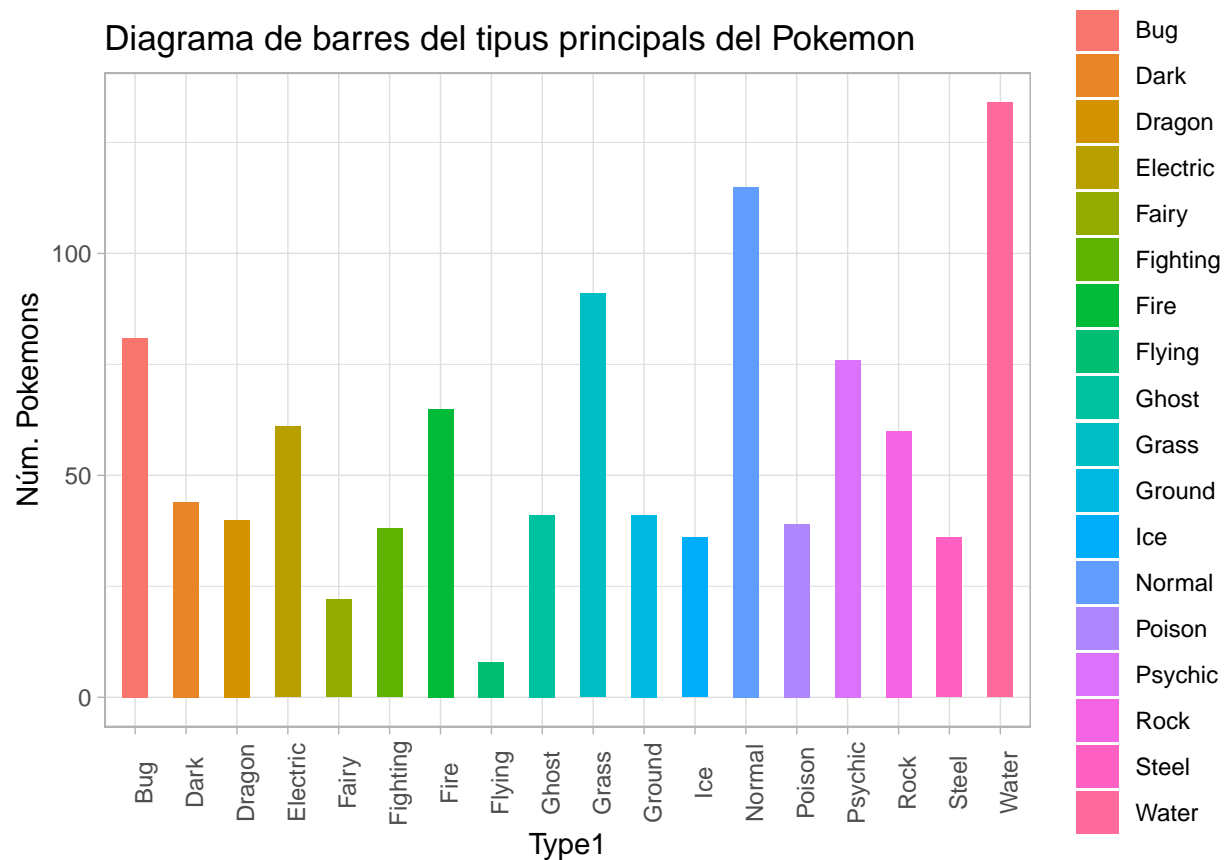
Anàlisi de la variables status



Es pot observar que la majoria dels Pokemons són normals, llavors de la classe Sub Legendary, Legendary i Mythical.

```
gg <- ggplot(pokedex, aes(x = type_1, fill = type_1)) +
  geom_bar(width=0.5) + # bar chart
  theme_light() + # theme
  labs(x = "Type1", # this labels the x axis
       y = "Núm. Pokemons", # this labels the y axis
       title = "Diagrama de barres del tipus principals del Pokemon") # title of the plot
gg+ theme(axis.text.x = element_text(angle = 90))
```

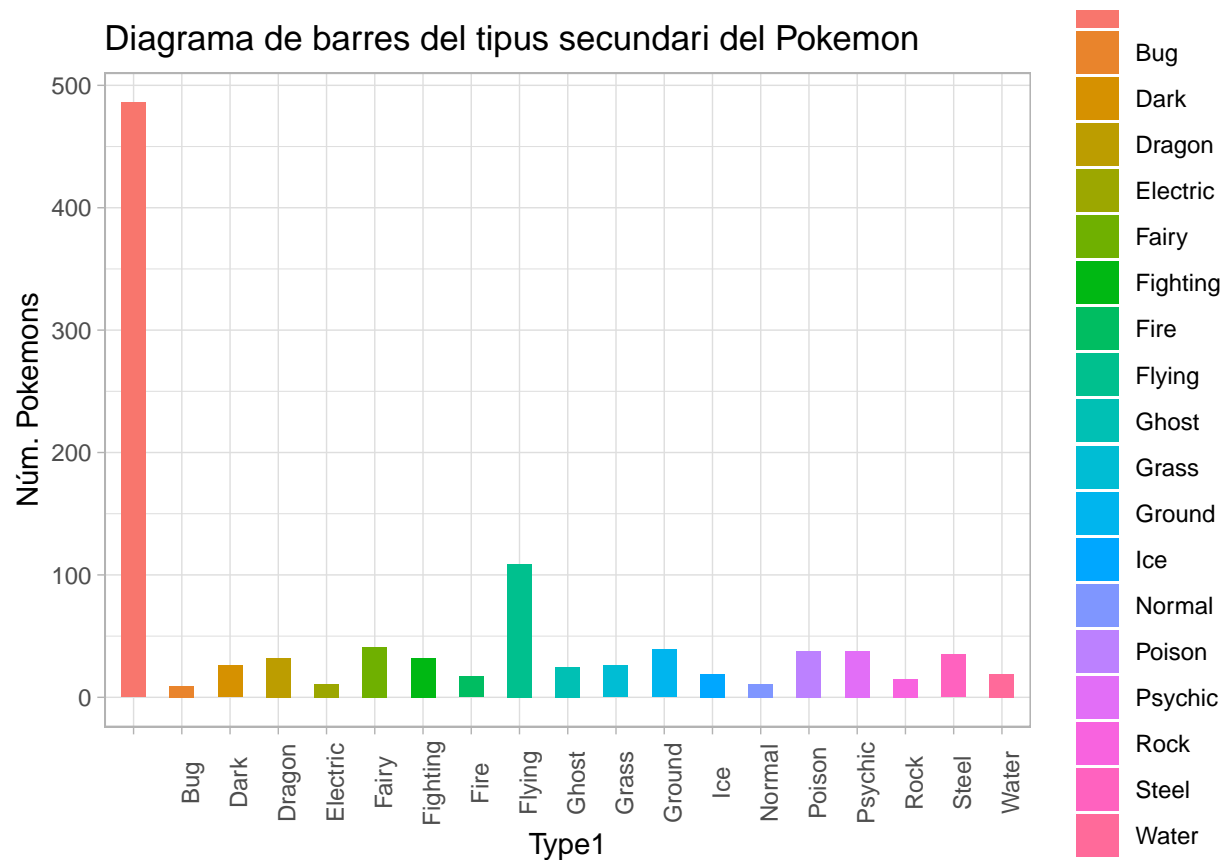
Anàlisi de la variable type\_1 (tipus principal)



Es pot observar que la majoria són de tipus water, normal, grass, bug i psychic.

```
gg <- ggplot(pokedex, aes(x = type_2, fill = type_2)) +
  geom_bar(width=0.5) + # bar chart
  theme_light() + # theme
  labs(x = "Type1", # this labels the x axis
       y = "Núm. Pokemons", # this labels the y axis
       title = "Diagrama de barres del tipus secundari del Pokemon") # title of the plot
gg+ theme(axis.text.x = element_text(angle = 90))
```

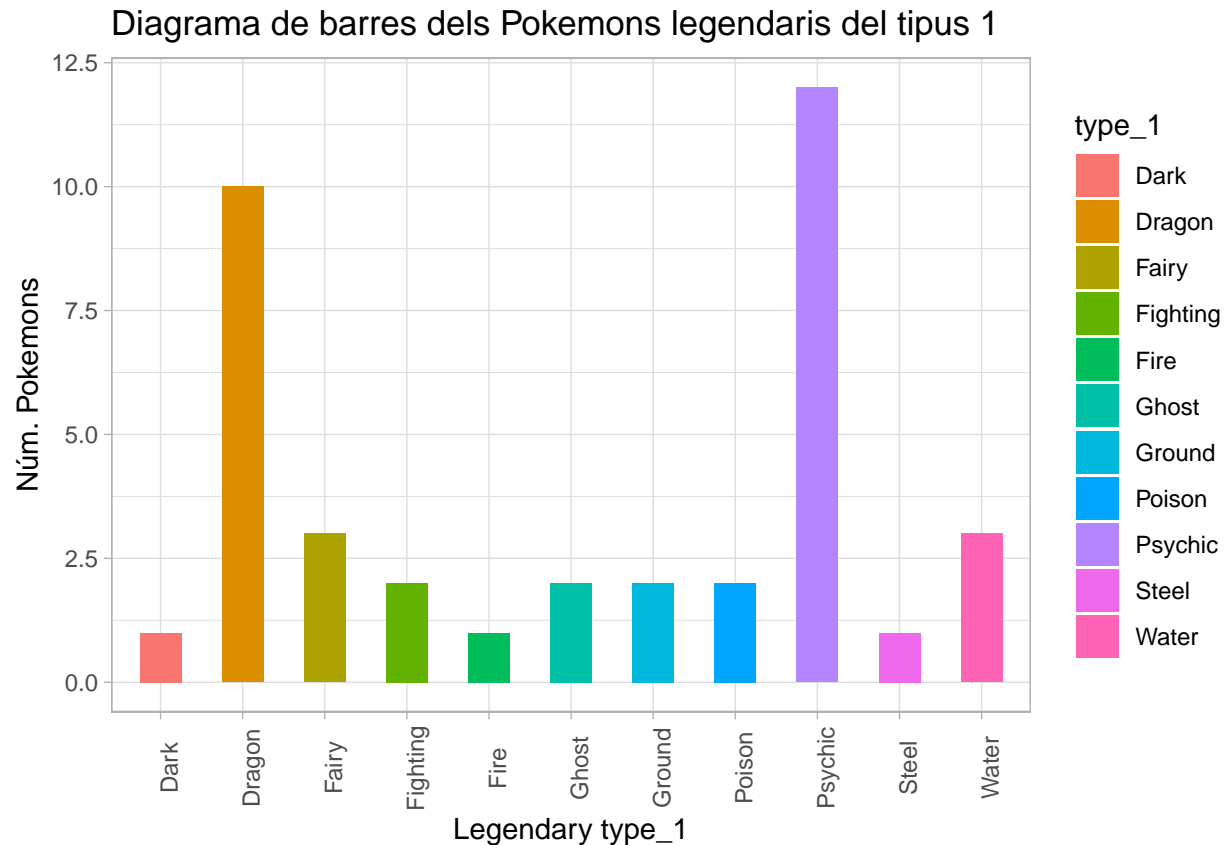
Anàlisi de la variable type\_2 (tipus secundari)



Es pot observar que la majoria no tenen tipus secundari, i si en tenen la majoria són de flying.

```
pokedex %>%
  dplyr::filter(status == "Legendary") %>%
  ggplot(aes(x = type_1, fill = type_1)) +
    geom_bar(width=0.5) + # bar chart
    theme_light() + # theme
    labs(x = "Legendary type_1", # this labels the x axis
         y = "Núm. Pokemons", # this labels the y axis
         title = "Diagrama de barres dels Pokemons legendaris del tipus 1")+ # title of the plot
    theme(axis.text.x = element_text(angle = 90))
```

Anàlisi dels tipus de Pokemons llegendaris

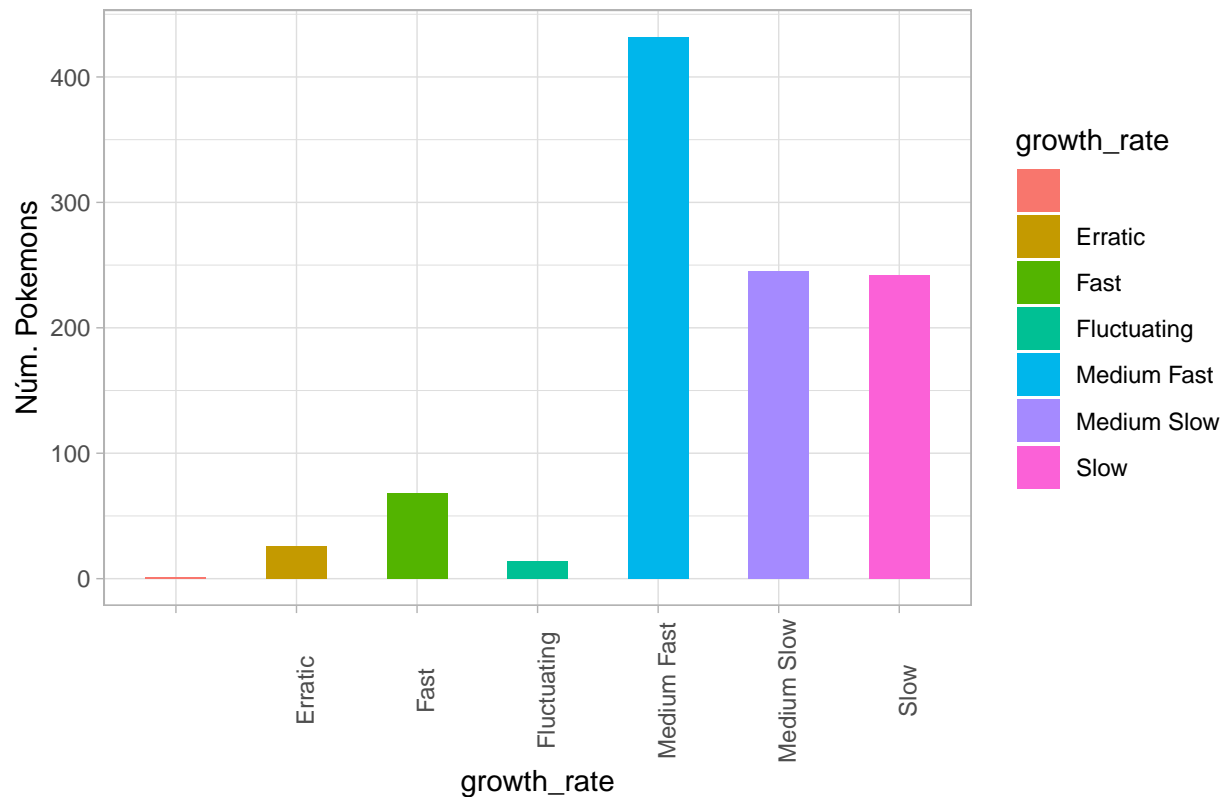


Es pot observar que la majoria de Pokemons llegendaris són de tipus Psychic i Dragon.

```
gg <- ggplot(pokedex, aes(x = growth_rate, fill = growth_rate)) +
  geom_bar(width=0.5) + # bar chart
  theme_light() + # theme
  labs(x = "growth_rate", # this labels the x axis
       y = "Núm. Pokemons", # this labels the y axis
       title = "Diagrama de barres del factor de creixement del Pokemon") # title of the plot
gg+ theme(axis.text.x = element_text(angle = 90))
```

Anàlisi de la variable growth\_rate (taxa de creixement)

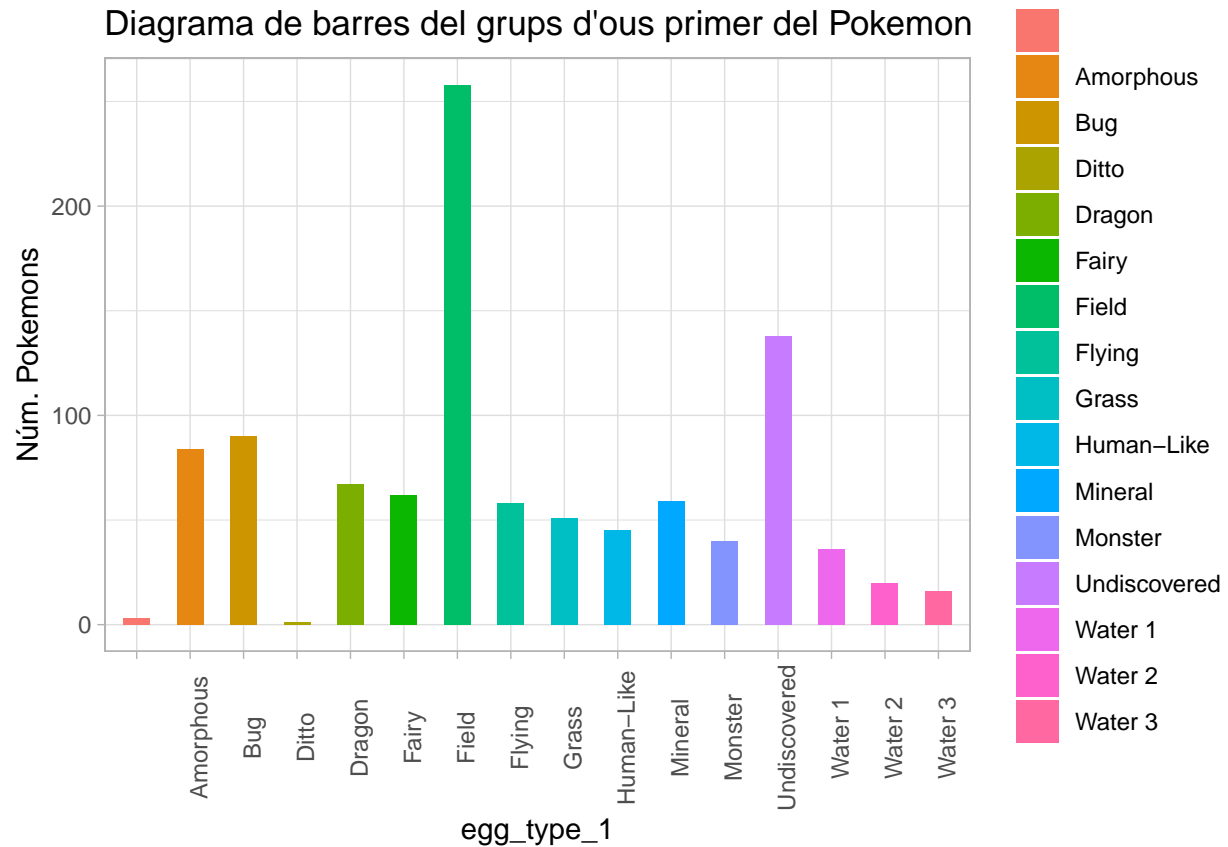
Diagrama de barres del factor de creixement del Pokemon



Es pot observar que la majoria creixen: Mediu Fast, Medium Slow i Slow.

```
gg <- ggplot(pokedex, aes(x = egg_type_1 , fill = egg_type_1 )) +
  geom_bar(width=0.5) + # bar chart
  theme_light() + # theme
  labs(x = "egg_type_1", # this labels the x axis
       y = "Núm. Pokemons", # this labels the y axis
       title = "Diagrama de barres del grups d'ous primer del Pokemon") # title of the plot
gg+ theme(axis.text.x = element_text(angle = 90))
```

Anàlisi de la variable egg\_type\_1 (Grups d'ous primer)

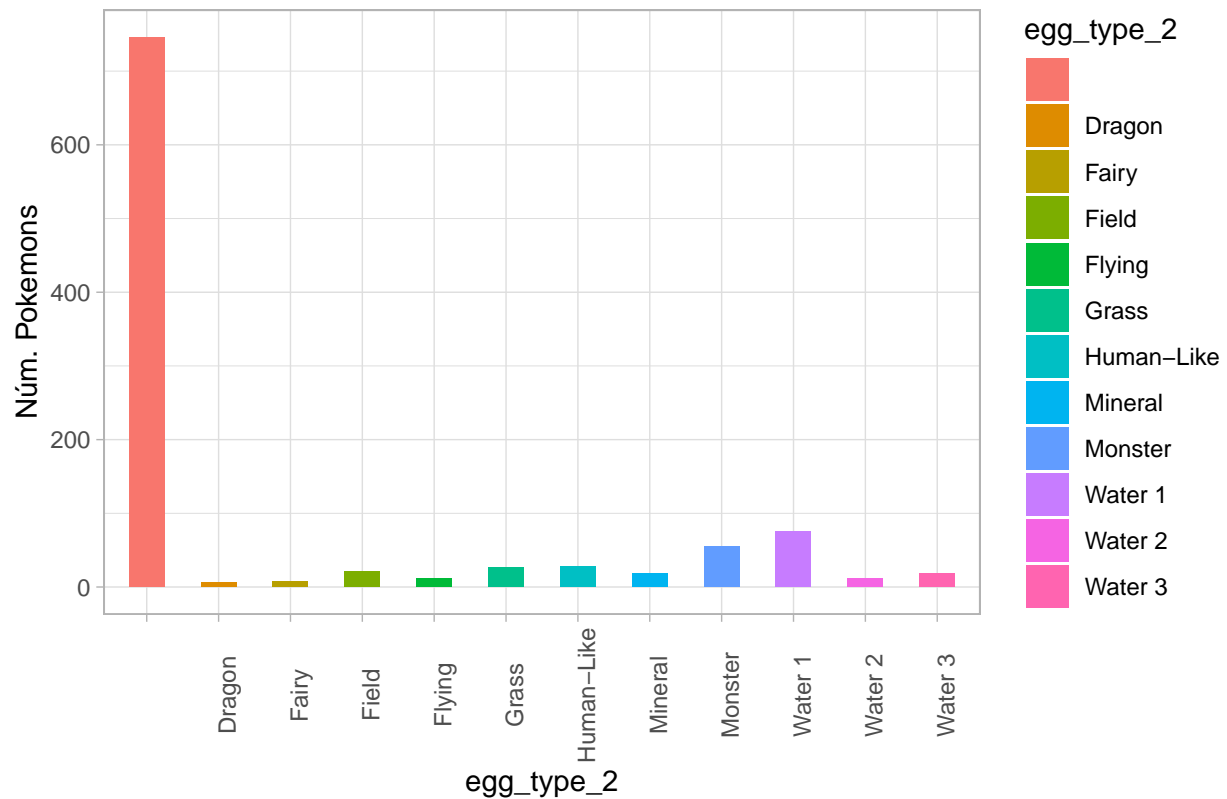


La majoria del tipus 1 eclosionen de tipus Field, Undiscovered, bug i Amorphous.

```
gg <- ggplot(pokedex, aes(x = egg_type_2 , fill = egg_type_2 )) +
  geom_bar(width=0.5) + # bar chart
  theme_light() + # theme
  labs(x = "egg_type_2", # this labels the x axis
       y = "Núm. Pokemons", # this labels the y axis
       title = "Diagrama de barres del grup d'ous segon del Pokemon") # title of the plot
gg+ theme(axis.text.x = element_text(angle = 90))
```

Anàlisi de la variable egg\_type\_2 (Grups d'ous segon)

Diagrama de barres del grup d'ous segon del Pokemon



La majoria del tipus 2 no tenen tipus, llavors són de tipus Water1 i Monster.

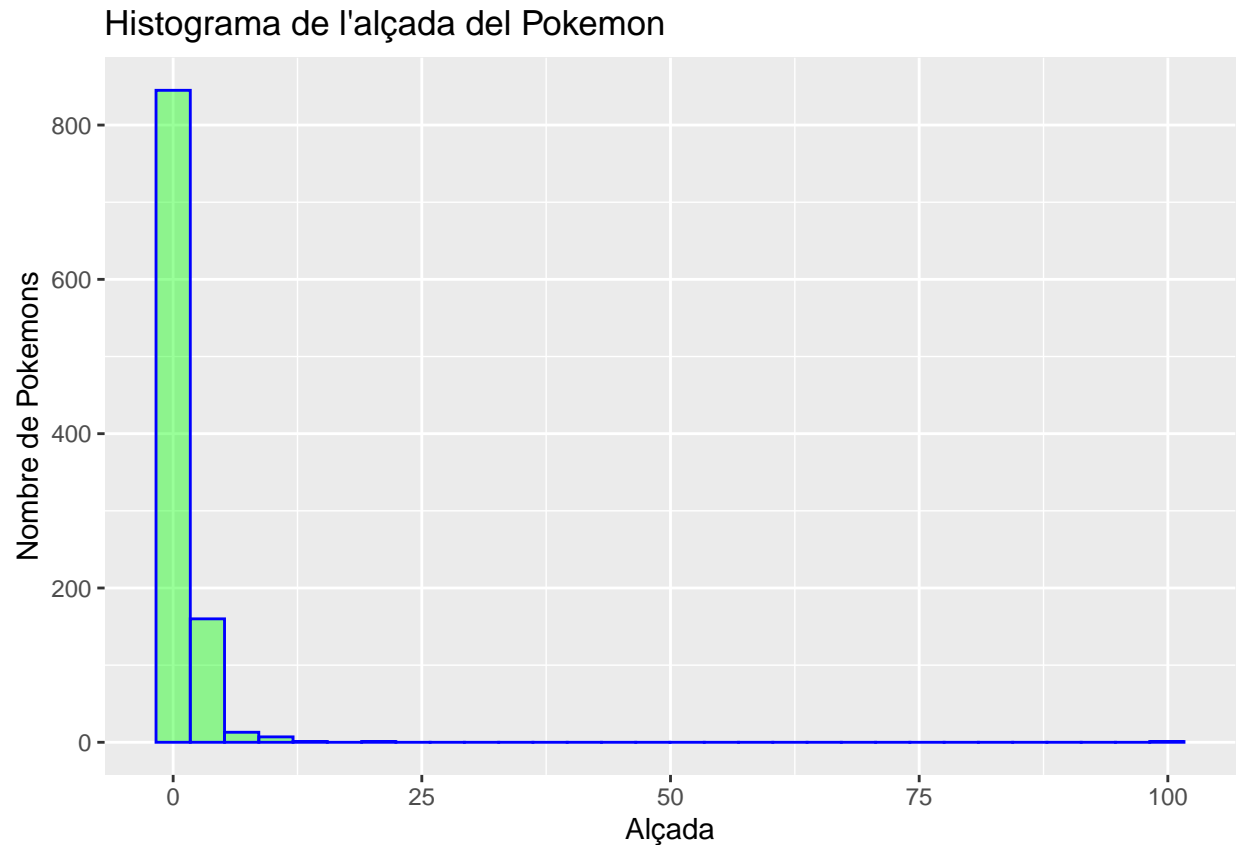
### Anàlisi de variables de tipus numèric informació de la pokedex

```
gg <- ggplot(data=pokedex, aes(height_m)) +
  geom_histogram( col="blue", fill="green", alpha = .4) +
  labs(title="Histograma de l'alçada del Pokemon") + labs(x="Alçada", y="Nombre de Pokemons")
gg
```

### Anàlisi de la variable height\_m

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```





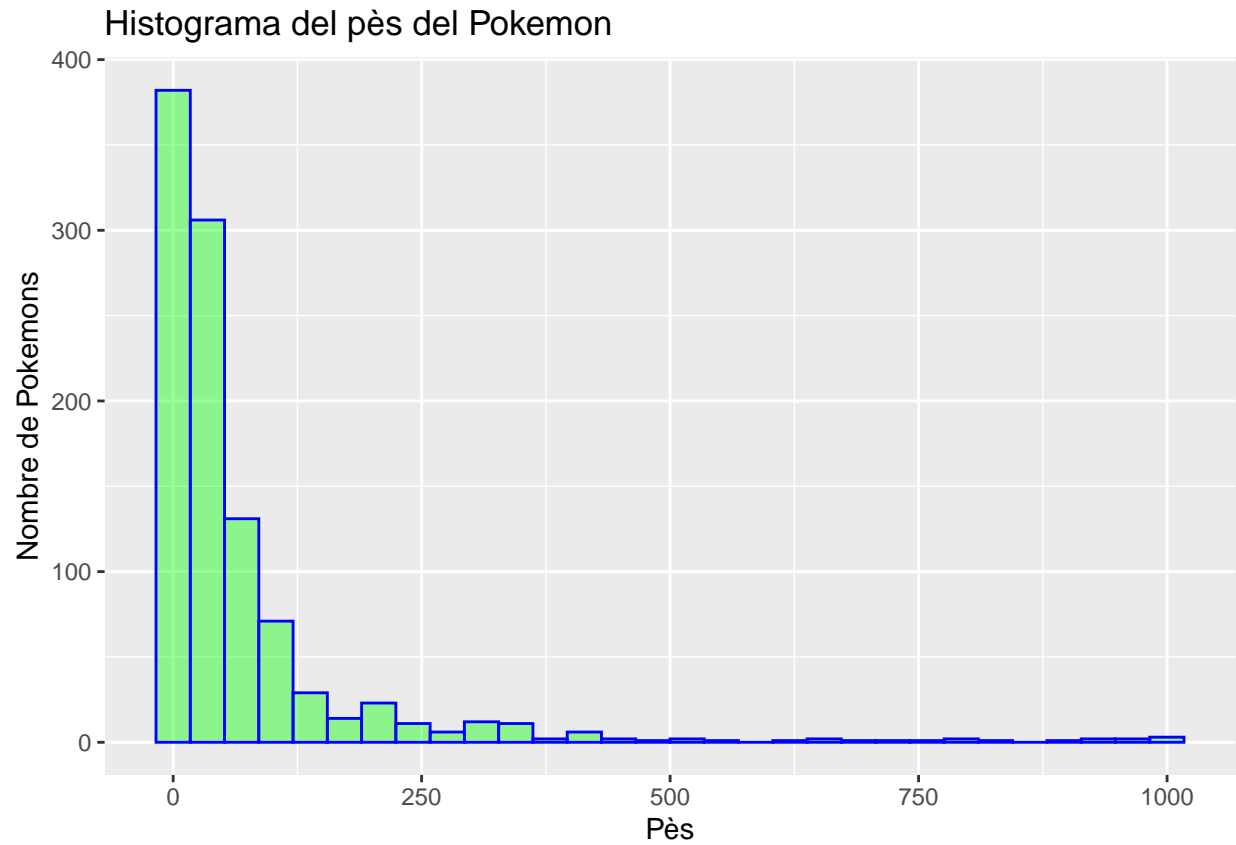
Es pot observar que la majoria de pokemons són d'estatura baixa de 0 a 1 metre.

```
gg <- ggplot(data=pokedex, aes(weight_kg)) +  
  geom_histogram( col="blue", fill="green", alpha = .4) +  
  labs(title="Histograma del pès del Pokemon") + labs(x="Pès", y="Nombre de Pokemons")  
gg
```

Anàlisi de la variable `weight_kg`

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
## Warning: Removed 1 rows containing non-finite values (stat_bin).
```

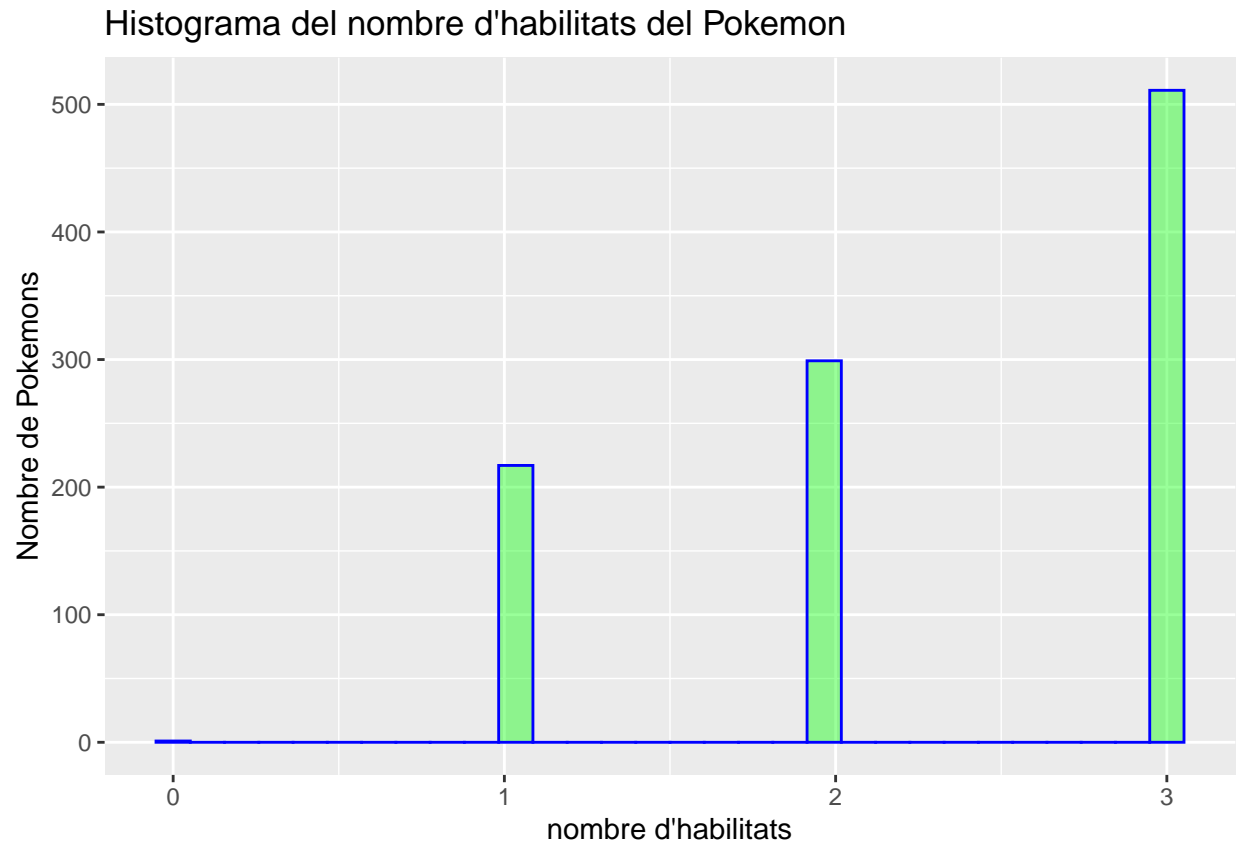


Es pot observar que la majoria de Pokemons, són lleugers de pocs Kg.

```
gg <- ggplot(data=pokedex, aes(abilities_number)) +
  geom_histogram( col="blue", fill="green", alpha = .4) +
  labs(title="Histograma del nombre d'habilitats del Pokemon") + labs(x="nombre d'habilitats", y="Nombre de Pokemons")
gg
```

**Anàlisi de la variable abilities\_number**

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Es pot observar que la majoria de Pokemons tenen 3 habilitats.

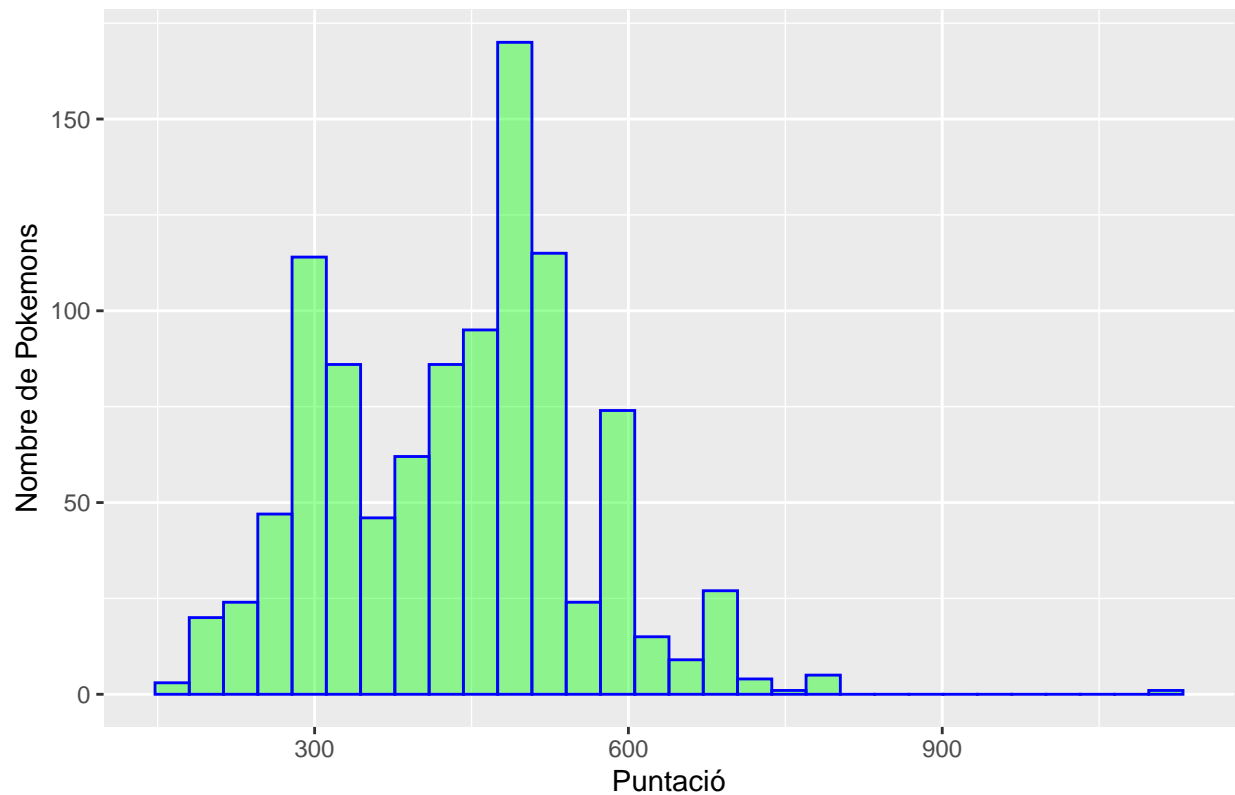
**Anàlisi de variables de tipus numèric informació de les estadístiques base**

```
gg <- ggplot(data=pokedex, aes(total_points)) +
  geom_histogram( col="blue", fill="green", alpha = .4) +
  labs(title="Histograma del nombre total de punts de base") + labs(x="Puntació", y="Nombre de Pokemons")
gg
```

**Anàlisi de la variable total-points**

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

Histograma del nombre total de punts de base



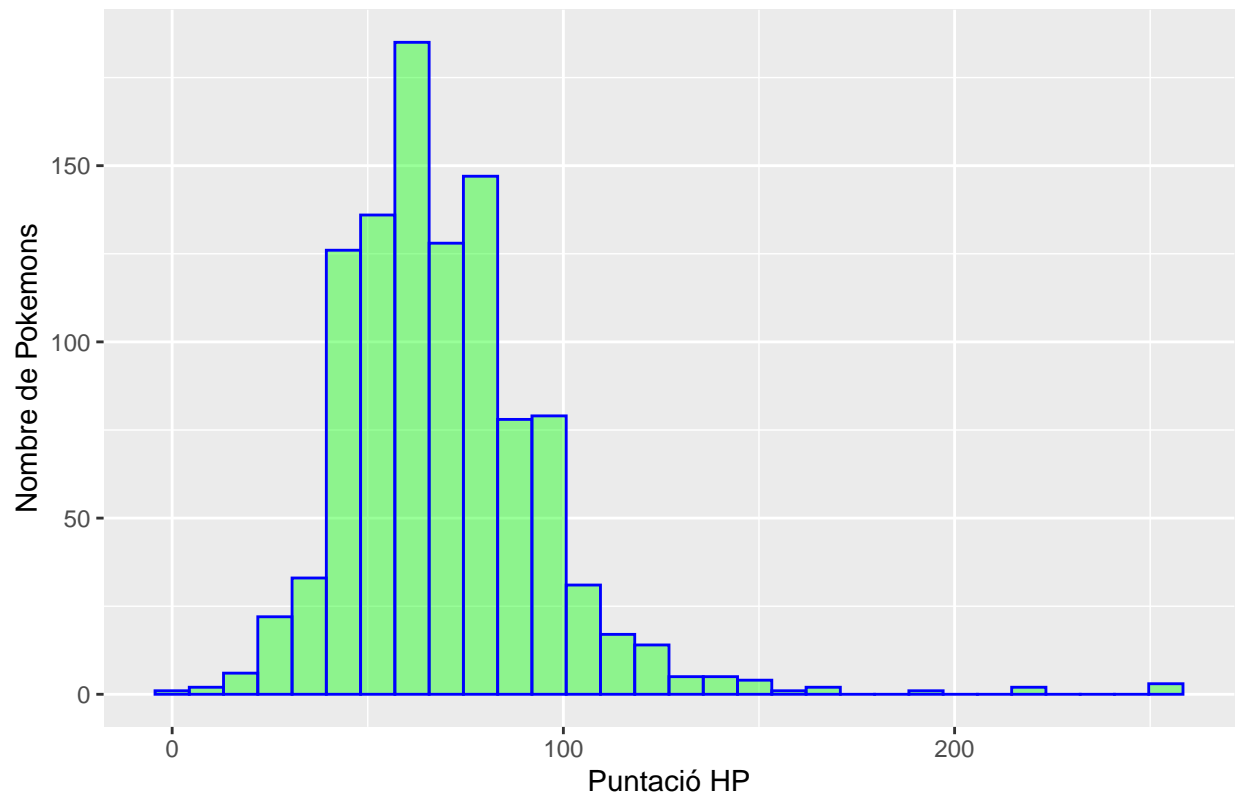
Es pot observar que la majoria del pokemons estan entre 300 i 600 punts totals.

```
gg <- ggplot(data=pokedex, aes(hp)) +  
  geom_histogram( col="blue", fill="green", alpha = .4) +  
  labs(title="Histograma del HP de base del Pokemon") + labs(x="Puntació HP", y="Nombre de Pokemons")  
gg
```

### Anàlisi de la variable hp

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

Histograma del HP de base del Pokemon



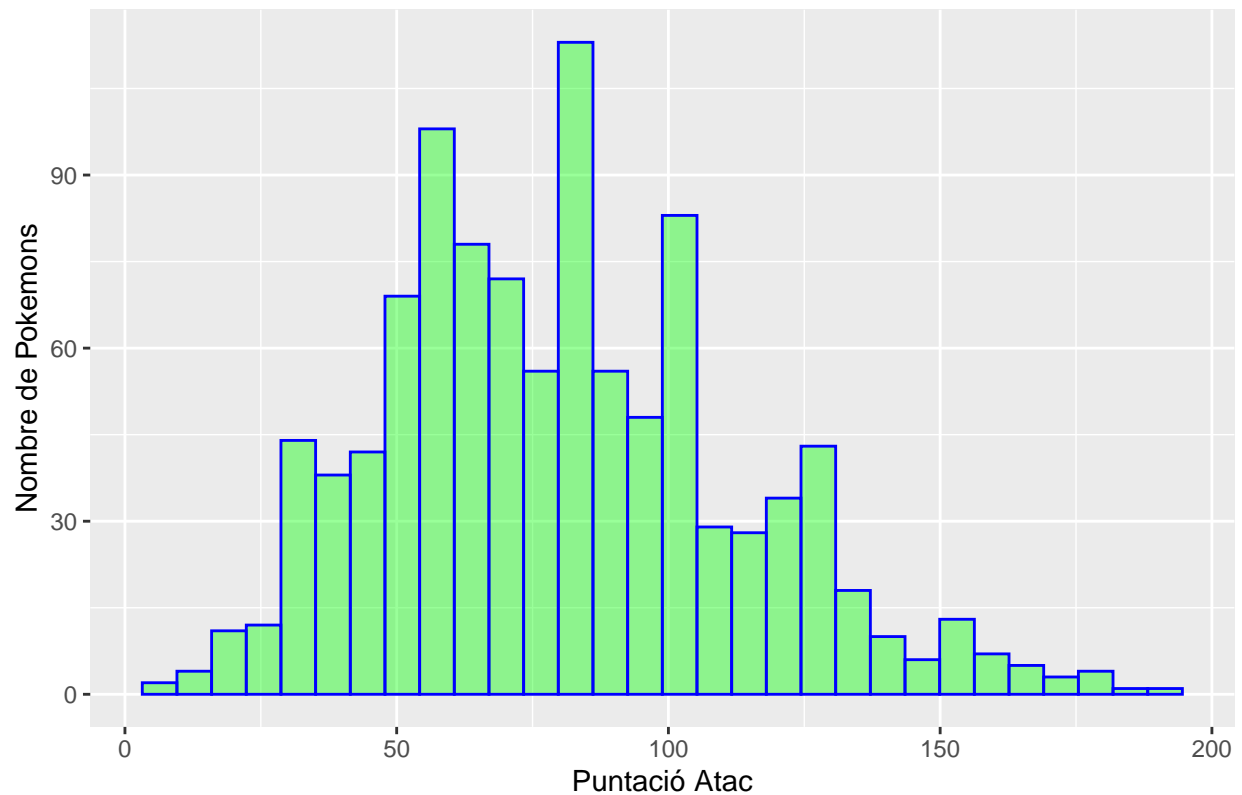
Es pot observar que la majoria estan entre 50 i 100 punts d'HP.

```
gg <- ggplot(data=pokedex, aes(attack)) +  
  geom_histogram( col="blue", fill="green", alpha = .4) +  
  labs(title="Histograma de l'atac de base del Pokemon") + labs(x="Puntació Atac", y="Nombre de Pok")  
gg
```

Anàlisi de la variable attack

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

Histograma de l'atac de base del Pokemon

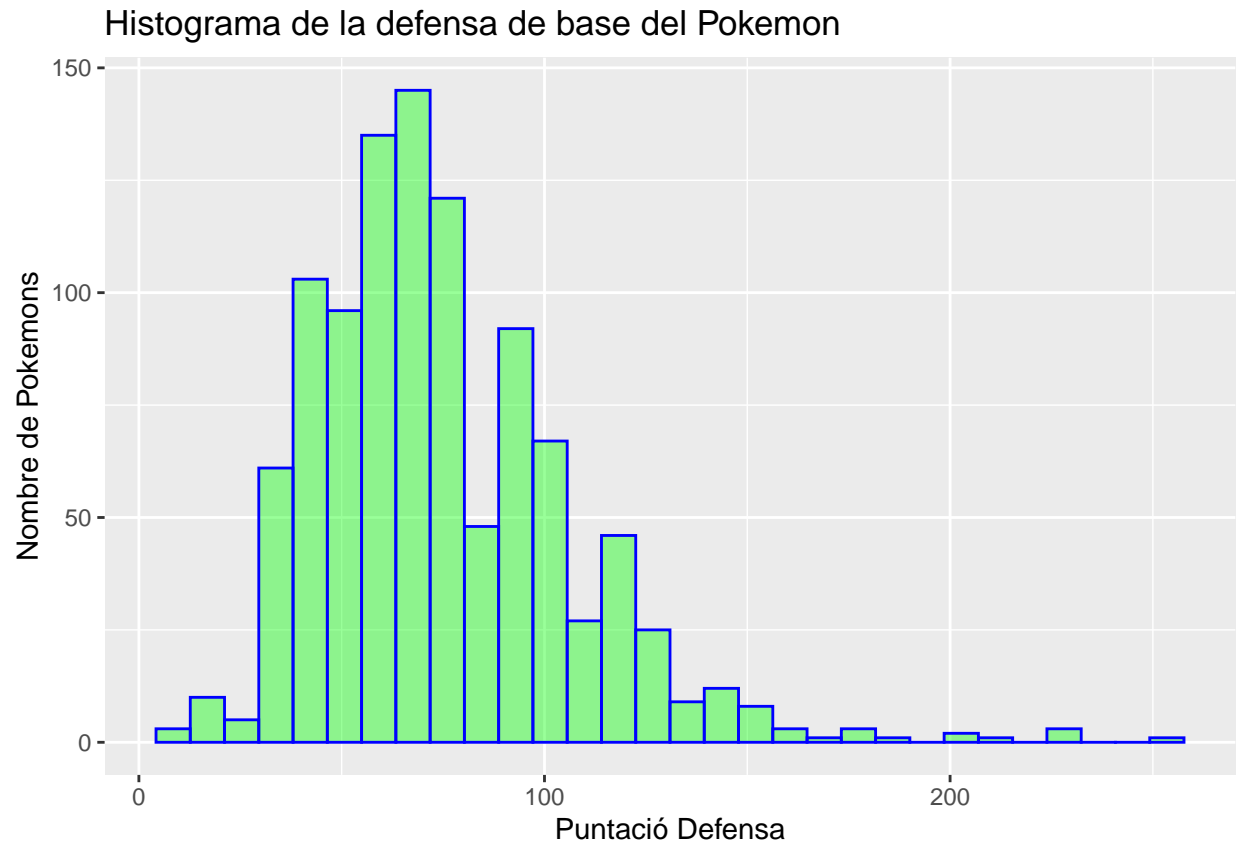


Es pot observar que la majoria estan entre 50 i 100 punts d'atac.

```
gg <- ggplot(data=pokedex, aes(defense)) +  
  geom_histogram( col="blue", fill="green", alpha = .4) +  
  labs(title="Histograma de la defensa de base del Pokemon") + labs(x="Puntació Defensa", y="Nombre  
gg
```

### Anàlisis de la variable defense

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



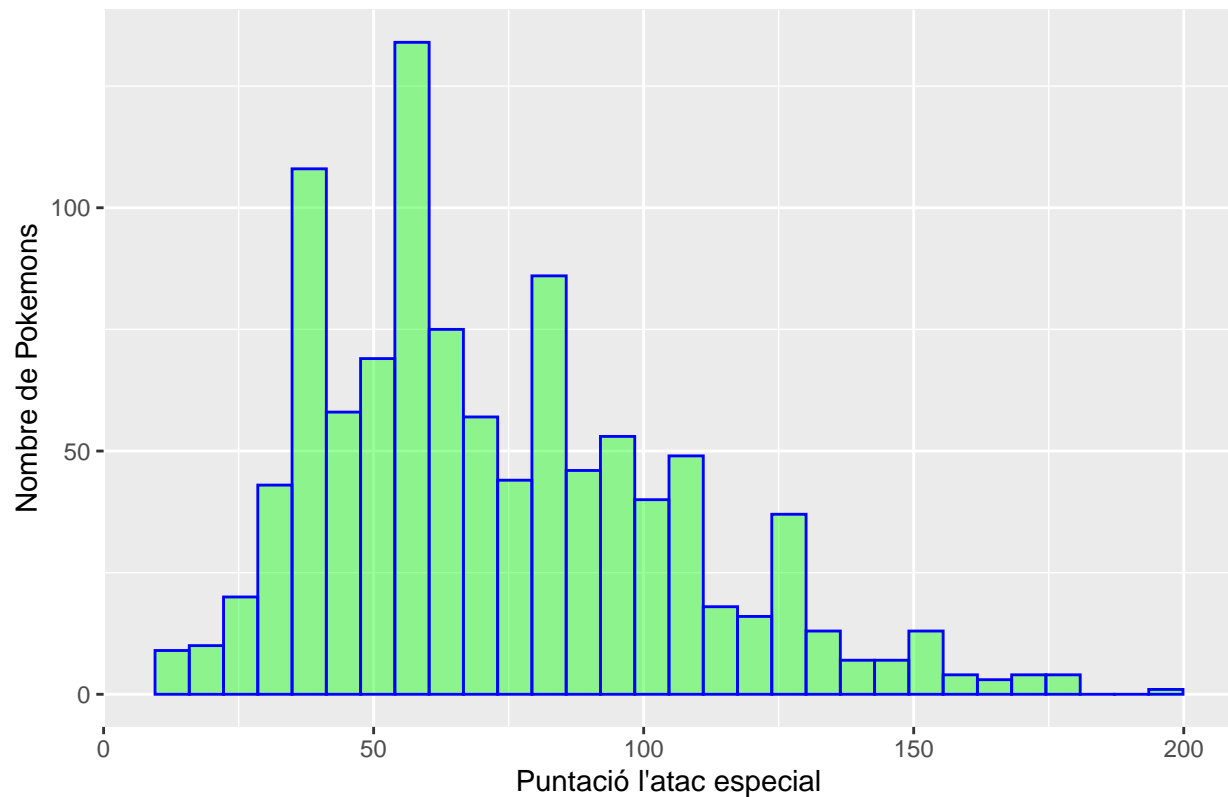
Es pot observar que la majoria estan entre 50 i 75 punts de defensa

```
gg <- ggplot(data=pokedex, aes(sp_attack)) +
  geom_histogram( col="blue", fill="green", alpha = .4) +
  labs(title="Histograma de l'atac especial de base del Pokemon") + labs(x="Puntació l'atac especial")
gg
```

**Anàlisi de la variable sp-attack**

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

Histograma de l'atac especial de base del Pokemon



La majoria estan entre 50 i 100 punts d'atac especial.

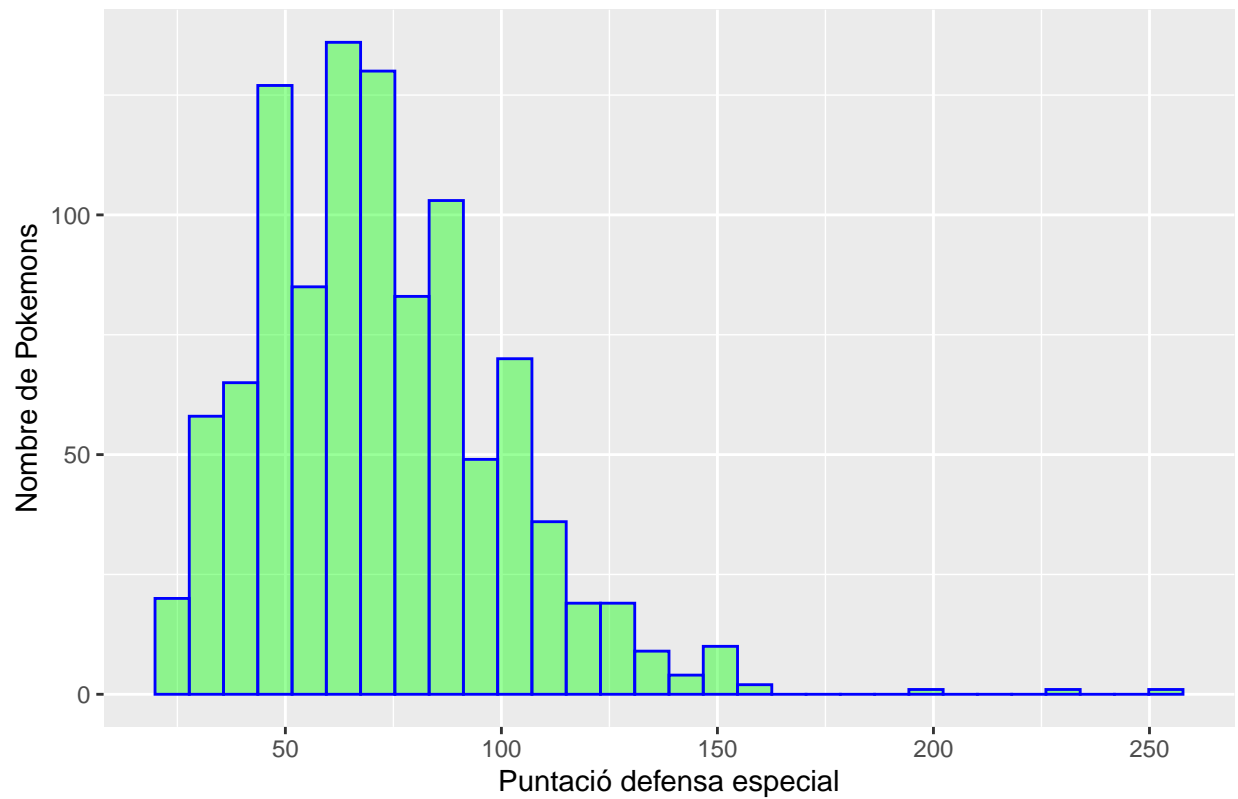
```
gg <- ggplot(data=pokedex, aes(sp_defense)) +
  geom_histogram( col="blue", fill="green", alpha = .4) +
  labs(title="Histograma de la defensa especial de base del Pokemon") + labs(x="Puntació defensa esp")
gg
```

Anàlisi de la variable sp-defense

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Histograma de la defensa especial de base del Pokemon



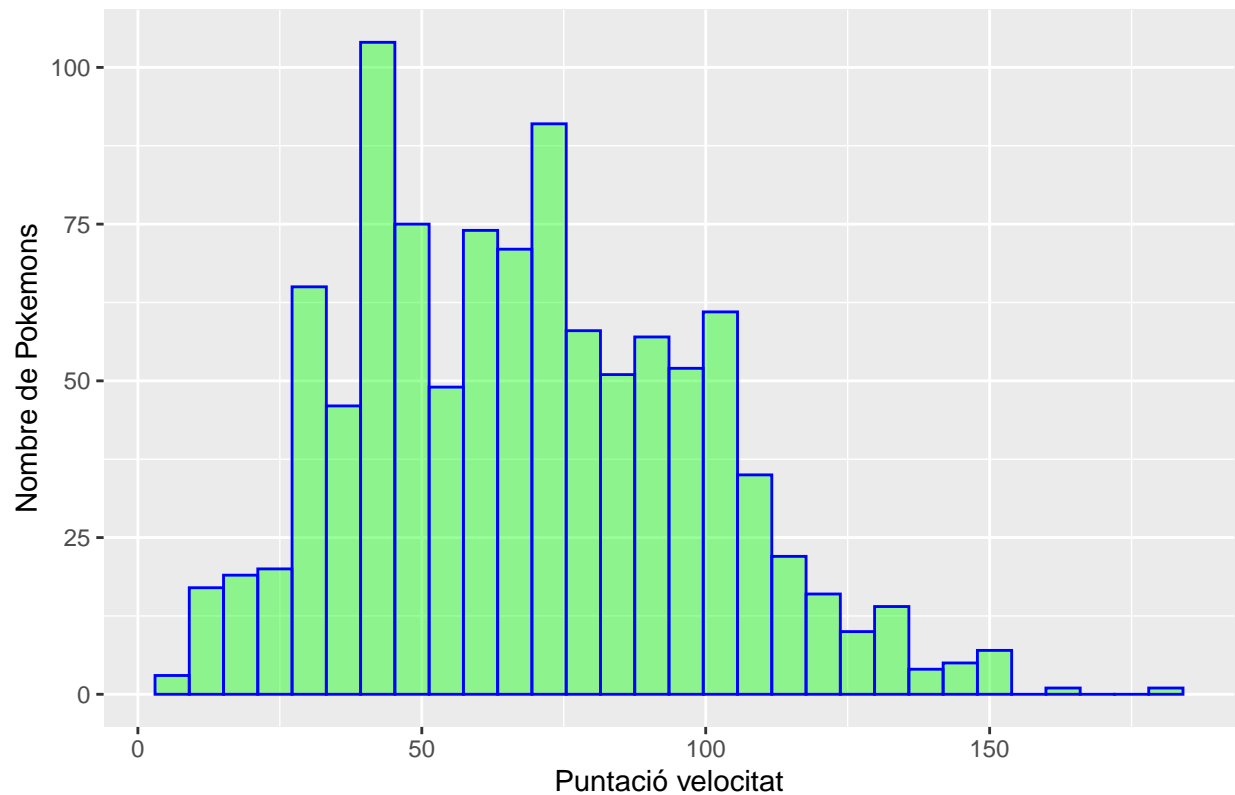
La majoria estan entre 50 i 90 punts de defensa especial.

```
gg <- ggplot(data=pokedex, aes(speed)) +
  geom_histogram( col="blue", fill="green", alpha = .4) +
  labs(title="Histograma de la velocitat de base del Pokemon") + labs(x="Puntació velocitat", y="Nombre de Pokemons")
gg
```

Anàlisi de la variable speed

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

Histograma de la velocitat de base del Pokemon

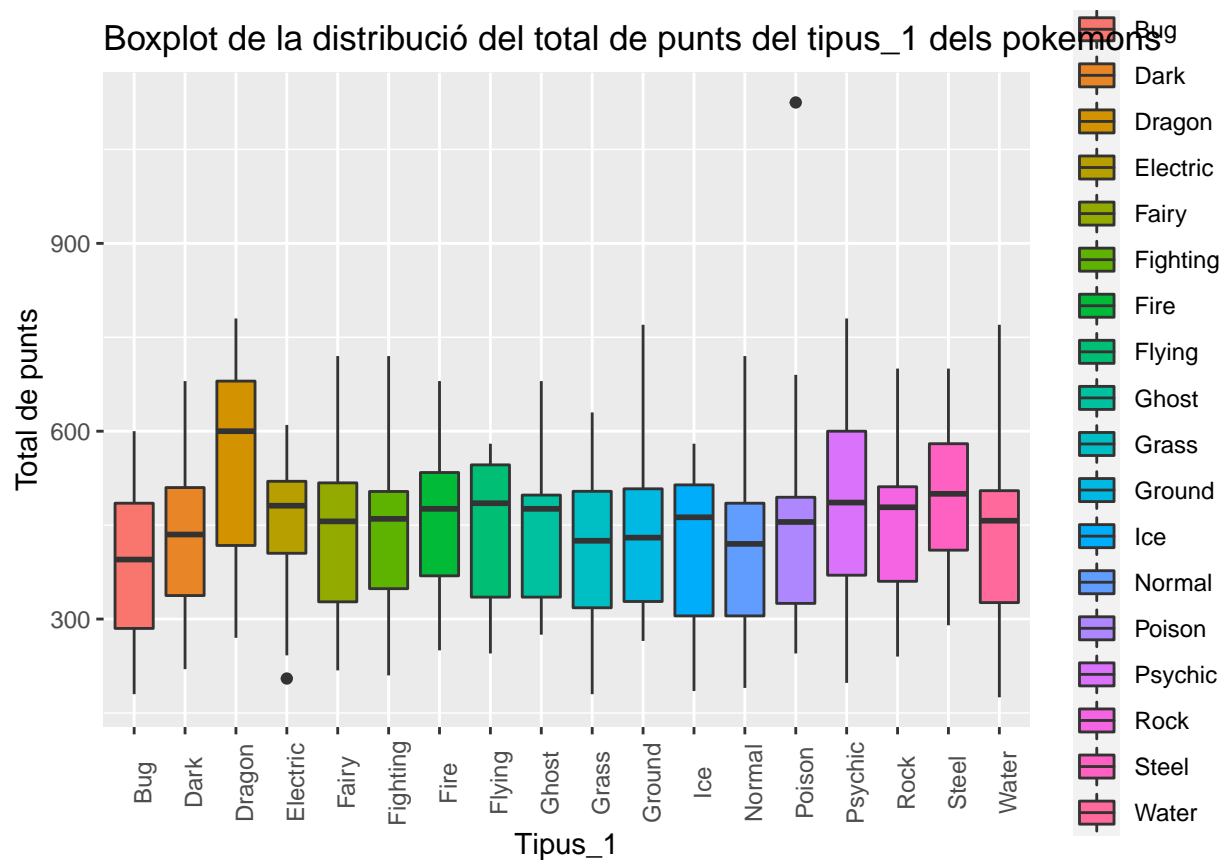


La majoria estan entre 25 i 100 punts de velocitat.

```
p <- ggplot(pokedex, aes(x=type_1, y=total_points, fill=type_1)) + geom_boxplot() +
  labs(title="Boxplot de la distribució del total de punts del tipus_1 dels pokemons") +
  labs(x="Tipus_1", y="Total de punts")

p + theme(axis.text.x = element_text(angle = 90))
```

Distribució del total de punts

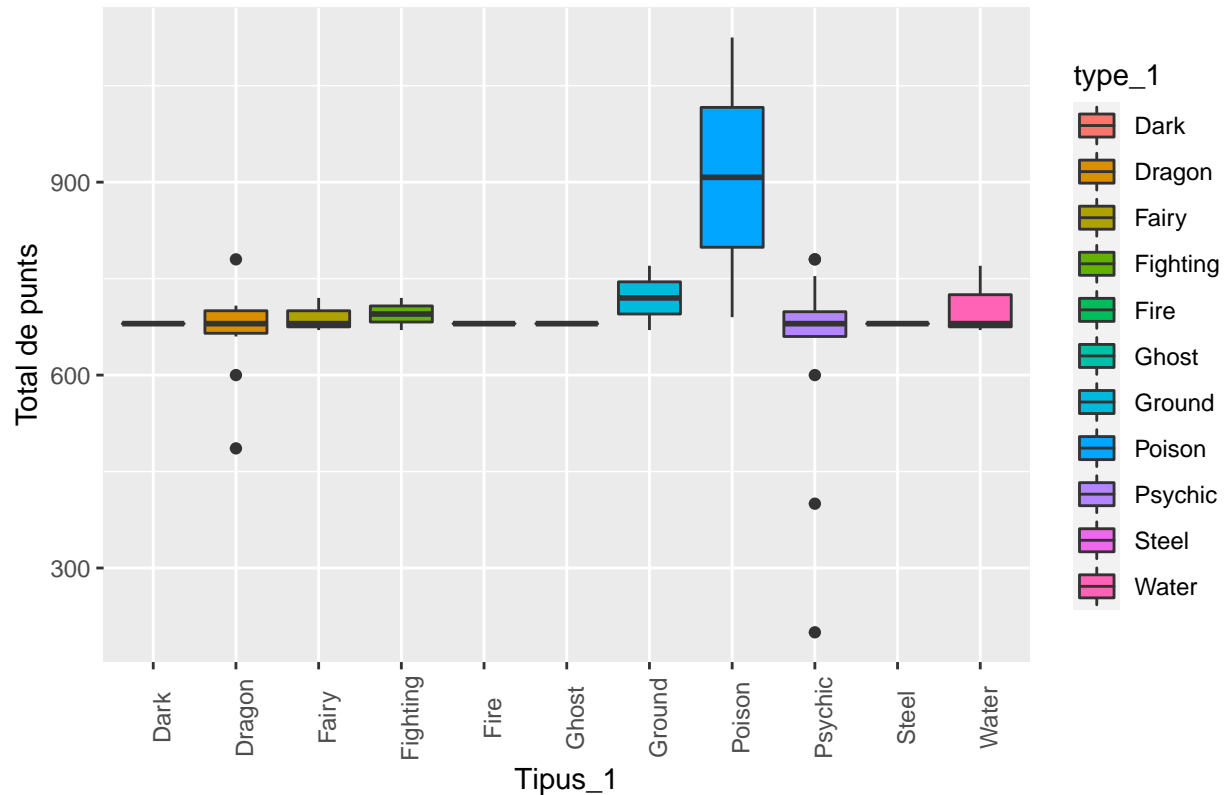


Es pot veure que de mitjana els pokemons que tenen més puntuació són els de tipus Dragon i llavors els de Psychic. Es pot veure que hi ha un valor absent a els de tipus Electric.

```
pokedex %>%
  dplyr::filter(status == "Legendary") %>%

  ggplot(aes(x=type_1, y=total_points, fill=type_1)) + geom_boxplot() +
  labs(title="Boxplot de la distribució del total de punts del tipus_1 dels pokemons llegendaris") +
  labs(x="Tipus_1", y="Total de punts") + theme(axis.text.x = element_text(angle = 90))
```

Boxplot de la distribució del total de punts del tipus\_1 dels pokemons lleger



Es pot observar que els que tenen més puntuació són els de tipus Poison. I llavors hi ha valors absents als de tipus Dragon i Psychic.

```
pokedexAnalysis <- subset(pokedex, select=c(status, type_number, type_1, type_2, height_m, weight_kg, a
```

```
head(pokedexAnalysis)
```

```
##   status type_number type_1 type_2 height_m weight_kg abilities_number
## 1 Normal           2 Grass Poison    0.7      6.9              2
## 2 Normal           2 Grass Poison    1.0     13.0              2
## 3 Normal           2 Grass Poison    2.0    100.0              2
## 4 Normal           2 Grass Poison    2.4   155.5              1
## 5 Normal           1  Fire          0.6      8.5              2
## 6 Normal           1  Fire          1.1     19.0              2
##   ability_1 ability_2 total_points hp attack defense sp_attack sp_defense speed
## 1 Overgrow              318 45    49    49      65      65    45
## 2 Overgrow              405 60    62    63      80      80    60
## 3 Overgrow              525 80    82    83     100     100    80
## 4 Thick Fat             625 80   100   123     122     120    80
## 5 Blaze                309 39    52    43      60      50    65
## 6 Blaze                405 58    64    58      80      65    80
##   catch_rate base_friendship base_experience growth_rate generation
## 1         45              70             64 Medium Slow          1
## 2         45              70            142 Medium Slow          1
## 3         45              70            236 Medium Slow          1
```

## 4	45	70	281 Medium Slow	1
## 5	45	70	62 Medium Slow	1
## 6	45	70	142 Medium Slow	1

### 3. Neteja de les dades.

#### 3.1. Les dades contenen zeros o elements buits? Com gestionaries aquests casos?

Les dades contenen zeros, però no estan considerades com a valors absents. Amb l'anàlisi del summary, es podia veure que hi ha atributs que tenen valors absents:

- weight\_kg: 1 NA
- catch\_rate: 104 NA
- base\_friendship: 104 NA
- base\_experience: 104 NA
- percentage\_male: 236 NA
- egg\_cycles: 1 NA

Amb la reducció del dataset:

```
sapply(pokedexAnalysis, function(x) sum(is.na(x)))
```

##	status	type_number	type_1	type_2
##	0	0	0	0
##	height_m	weight_kg	abilities_number	ability_1
##	0	1	0	0
##	ability_2	total_points	hp	attack
##	0	0	0	0
##	defense	sp_attack	sp_defense	speed
##	0	0	0	0
##	catch_rate	base_friendship	base_experience	growth_rate
##	104	104	104	0
##	generation			
##	0			

En el cas de weight\_kg, el substituïrem a partir dels k-veïns més propers utilitzant la distància de Grower amb la informació de totes les variables. Els altres atributs absents, no els analitzarem. Ja que ens centrem l'anàlisi en altres atributs.

```
library(VIM)
```

```
## Loading required package: colorspace
```

```
## Loading required package: grid
```

```
## Loading required package: data.table
```

```
## VIM is ready to use.
```

```
## Since version 4.0.0 the GUI is in its own package VIMGUI.
```

```
##
```

```
## Please use the package to use the new (and old) GUI.
```

```
## Suggestions and bug-reports can be submitted at: https://github.com/alexxkova/VIM/issues
```

```
##
```

```
## Attaching package: 'VIM'
```

```
## The following object is masked from 'package:datasets':
```

```
##
```

```
##      sleep
```

```
pokedexKNN <- kNN(pokedexAnalysis)
pokedexAnalysis$weight_kg <- pokedexKNN$weight_kg
```

```
### verifiquem que no te valors perduts
sum(is.na(pokedex$weight_kg))
```

```
## [1] 1
```

### 3.2. Identificació i tractament de valors extrems.

Els valors atípics o extrems, són les observacions que són numèricament distants de la resta de les observacions. A vegades poden indicar que aquests valors pertanyen a una població diferent de les mostres establertes.

A continuació analitzem els valors atípics:

```
for(k in 1:ncol(pokedexAnalysis)){
  if(is.numeric(pokedexAnalysis[,k])){
    x <- boxplot.stats(pokedexAnalysis[,k])$out
    print(colnames(pokedexAnalysis[k]))
    print(x)
  }
}
```

```
## [1] "type_number"
```

```
## integer(0)
```

```
## [1] "height_m"
```

```
## [1] 3.5 8.8 10.9 3.0 6.5 6.5 4.0 9.2 10.5 5.2 3.8 14.5
```

```
## [13] 6.2 4.5 9.8 3.5 5.0 7.0 10.8 5.4 4.2 3.7 4.5 6.9
```

```
## [25] 3.2 3.3 3.0 3.2 2.9 3.0 3.3 3.6 3.0 5.8 5.0 4.5
```

```
## [37] 6.5 8.2 3.0 3.9 3.4 4.0 3.8 9.2 5.5 3.8 4.2 7.5
```

```
## [49] 3.6 5.5 3.8 3.0 3.0 3.0 3.0 2.9 2.9 20.0 100.0
```

```
## [1] "weight_kg"
```

```
## [1] 300.0 316.0 210.0 415.6 235.0 305.0 220.0 460.0 210.0 400.0 740.0 220.0
```

```
## [13] 178.0 198.0 187.0 202.0 255.0 216.0 199.0 253.8 360.0 395.0 398.0 220.0
```

```
## [25] 320.5 168.0 162.0 256.5 350.2 202.5 550.0 942.9 230.0 175.0 205.0 352.0
```

```
## [37] 430.0 950.0 999.7 206.5 392.0 310.0 187.0 300.0 185.0 180.0 282.8 291.0
```

```
## [49] 340.0 683.0 336.0 430.0 420.0 750.0 650.0 320.0 260.0 200.5 200.0 260.0
```

```
## [61] 330.0 250.0 260.0 200.0 330.0 345.0 325.0 325.0 325.0 270.0 225.0 505.0
```

```
## [73] 215.0 203.0 305.0 610.0 490.0 195.0 180.0 920.0 250.0 212.0 185.0 210.0
```

```
## [85] 999.9 230.0 333.6 999.9 888.0 230.0 460.0 350.0 230.0 820.0 800.0 310.5
```

```
## [97] 520.0 650.0 190.0 215.0 175.0 355.0 785.0 210.0 950.0 355.0
```

```
## [1] "abilities_number"
```

```

## [1] 0
## [1] "total_points"
## [1] 1125
## [1] "hp"
## [1] 140 250 130 130 160 190 255 150 1 144 130 170 150 135 150 150 165 126 126
## [20] 216 137 137 223 135 140 255
## [1] "attack"
## [1] 190 185 180 180 180 170 170 181 170
## [1] "defense"
## [1] 180 180 160 200 230 230 180 230 200 160 160 168 184 152 211 250
## [1] "sp_attack"
## [1] 175 170 194 165 165 180 180 180 170 170 173 167
## [1] "sp_defense"
## [1] 230 154 154 200 160 160 154 250
## [1] "speed"
## [1] 160 180
## [1] "catch_rate"
## [1] 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255
## [20] 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255
## [39] 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255
## [58] 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255 255
## [1] "base_friendship"
## [1] 140 140 140 35 35 35 35 35 35 0 0 0 100 140 35 35 35 35
## [19] 35 35 35 140 35 35 35 35 35 35 35 0 0 100 35 35 35 35
## [37] 35 35 35 35 35 35 35 35 35 35 35 35 35 35 35 35 35 35
## [55] 35 35 35 35 35 35 35 35 35 35 35 90 90 90 90 0 0 0
## [73] 0 0 0 100 0 0 0 0 100 100 100 0 140 140 35 35 140 35
## [91] 100 35 35 35 35 35 35 140 140 140 0 0 100 0 0 0 100 0
## [109] 100 100 0 100 35 35 35 35 35 35 35 35 35 35 35 35 35 35
## [127] 90 90 90 90 0 0 90 90 0 0 0 35 35 100 100 0 35 35
## [145] 35 0 0 0 0 0 0 100 100 100 0 0 0 0 0 0 0 0
## [163] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## [1] "base_experience"
## [1] 608 425
## [1] "generation"
## integer(0)

```

No tractarem aquests valors extrems, degut a la tipologia de problema que estem tractant, com que hi ha tanta varietat d'espècies de Pokemon, aquests valors poden ser normals.

## 4. Anàlisi de les dades.

### 4.1. Selecció dels grups de dades que es volen analitzar/comparar (planificació dels anàlisis a aplicar).

Pels diferents anàlisis que es vol realitzar, es pretén analitzar els Pokemons per grups de classe separats: Tot el grup en general, per el seu estatus: normal, llegendari i sub llegendari i per tipus de pokemon: Bug, Dragon, Psychic, Steel i Poison. Aquests tipus de Pokemons, han estat seleccionats segons el que hem vist en els gràfics dels boxplots de distribució de total de punts i tipus\_1. I de la variable status. Per tant haurem de crear diferents dataframes per poder fer els estudis corresponents.

Es generaran variables de tipus Factor tant de pes com d'alçada. Pes[lleuger,pesat] i Alçada[baix,alt], per poder realitzar test de 2 mostres amb el total\_points.

```

funcH <- function(h) {
  if(h <= 1.368 ){
    "baix"
  }else if(h > 1.368 ){
    "alt"
  }
}

funcW <- function(w) {
  if(w <= 69.75 ){
    "lleuger"
  }else if(w > 69.75 ){
    "pesat"
  }
}

##"height_m", "weight_kg"
pokedexAnalysis$Fheight_m <- as.factor(sapply(pokedexAnalysis$height_m,funcH))
pokedexAnalysis$Fweight_kg <- as.factor(sapply(pokedexAnalysis$weight_kg,funcW))

summary(pokedexAnalysis)

```

```

##           status      type_number      type_1      type_2
## Legendary      : 39   Min.      :1.000   Water  :134           :486
## Mythical       : 29   1st Qu.:1.000   Normal  :115   Flying  :109
## Normal         :915   Median :2.000   Grass   : 91   Fairy   : 41
## Sub Legendary: 45   Mean    :1.527   Bug     : 81   Ground  : 39
##               3rd Qu.:2.000   Psychic: 76   Poison  : 38
##               Max.    :2.000   Fire    : 65   Psychic: 38
##               (Other):466   (Other):277
##      height_m      weight_kg      abilities_number      ability_1
## Min.      : 0.100   Min.      : 0.10   Min.      :0.000   Levitate   : 41
## 1st Qu.: 0.600   1st Qu.: 8.80   1st Qu.:2.000   Swift Swim : 30
## Median : 1.000   Median : 28.65   Median :2.000   Keen Eye   : 28
## Mean    : 1.368   Mean    : 70.03   Mean     :2.284   Chlorophyll: 25
## 3rd Qu.: 1.500   3rd Qu.: 70.00   3rd Qu.:3.000   Pressure   : 25
## Max.    :100.000   Max.    :999.90   Max.      :3.000   Blaze      : 24
##               (Other) :855
##      ability_2      total_points      hp      attack
##               :515   Min.      : 175.0   Min.      : 1.00   Min.      : 5.00
## Frisk           : 17   1st Qu.: 330.0   1st Qu.: 50.00   1st Qu.: 55.00
## Sturdy          : 14   Median : 455.0   Median : 66.50   Median : 76.00
## Shell Armor: 13   Mean    : 437.6   Mean     : 69.58   Mean     : 80.12
## Early Bird : 10   3rd Qu.: 510.0   3rd Qu.: 80.00   3rd Qu.:100.00
## Gluttony      : 10   Max.    :1125.0   Max.      :255.00   Max.      :190.00
## (Other)       :449
##      defense      sp_attack      sp_defense      speed
## Min.      : 5.00   Min.      : 10.00   Min.      : 20.00   Min.      : 5.00
## 1st Qu.: 50.00   1st Qu.: 50.00   1st Qu.: 50.00   1st Qu.: 45.00
## Median : 70.00   Median : 65.00   Median : 70.00   Median : 65.00
## Mean    : 74.48   Mean     : 72.73   Mean     : 72.13   Mean     : 68.53
## 3rd Qu.: 90.00   3rd Qu.: 95.00   3rd Qu.: 90.00   3rd Qu.: 90.00
## Max.    :250.00   Max.      :194.00   Max.      :250.00   Max.      :180.00

```



```
##
##   catch_rate    base_friendship  base_experience    growth_rate
##   Min.       : 3.00    Min.       : 0.00    Min.       : 36.0           : 1
##   1st Qu.: 45.00    1st Qu.: 70.00    1st Qu.: 67.0    Erratic       : 26
##   Median : 60.00    Median : 70.00    Median :159.0    Fast          : 68
##   Mean    : 93.17    Mean    : 64.14    Mean    :153.8    Fluctuating: 14
##   3rd Qu.:127.00    3rd Qu.: 70.00    3rd Qu.:201.5    Medium Fast:432
##   Max.     :255.00    Max.     :140.00    Max.     :608.0    Medium Slow:245
##   NA's     :104     NA's     :104     NA's     :104     Slow          :242
##   generation    Fheight_m      Fweight_kg
##   Min.       :1.000    alt :340    lleuger:770
##   1st Qu.:2.000    baix:688    pesat  :258
##   Median :4.000
##   Mean      :4.034
##   3rd Qu.:6.000
##   Max.      :8.000
##
```

A continuació, generem els diferents dataframes, a partir del dataframe que prèviament s'ha netejat "pokedex-Analysis":

```
pokedexAnalysis.Legendary <-subset(pokedexAnalysis, status == 'Legendary')
pokedexAnalysis.SubLegendary <-subset(pokedexAnalysis, status == 'Sub Legendary')
pokedexAnalysis.Normal <-subset(pokedexAnalysis, status == 'Normal')
pokedexAnalysis.Bug <-subset(pokedexAnalysis, type_1 == 'Bug')
pokedexAnalysis.Dragon <-subset(pokedexAnalysis, type_1 == 'Dragon')
pokedexAnalysis.Psychic <-subset(pokedexAnalysis, type_1 == 'Psychic')
pokedexAnalysis.Poison <-subset(pokedexAnalysis, type_1 == 'Poison')
pokedexAnalysis.Steel <-subset(pokedexAnalysis, type_1 == 'Steel')
```

Els estudis a realitzar seran:

1. Estudiar les variables d'estadística base, que fan determinar la força total d'un Pokemon. Es vol estudiar la relació entre aquests atributs. S'estudiarà de forma global entre tots els tipus i llavors amb la classe Dragon i Bug, que són els més forts i debils respectivament. També estudiarem les classes Steel, Poison, Psychic. Influeix igual en els Pokemons llegendaris i sub llegendaris?
2. Estudiar si segons diferents tipus de pes i d'alçada del Pokemon, influeixen a que el Pokemon sigui més fort, per tant tinguin un total de punts més elevat.
3. Fer una predicció de com les variables d'estadística base, el pes i l'alçada influeixen amb la variable base\_experience.

#### 4.2. Comprovació de la normalitat i homogeneïtat de la variància.

Mitjançant la prova de normalitat *Shapiro-Wilk* comprovem la normalitat de les dades quantitatives. Si el p-valor és superior a 0,05 ho considerarem com distribució normal:

```
## definim taula de variables per els càlculs de la normalitat
analisi_normalitat <- c("height_m", "weight_kg","total_points", "hp", "attack", "defense", "sp_attack",
```

```
## funció per calcular el test de Shapiro-Wilk
shapirotest <- function(x){
  alpha = 0.05
  p_val = shapiro.test(x)$p.value
  if (p_val > alpha) { "> 0.05" }
  else {"< 0.05"}
}
```

```
apply(pokedexAnalysis[,analisi_normalitat], 2, shapirotest)
```

Comprovació de la normalitat de tot el grup

```
##      height_m      weight_kg total_points      hp      attack      defense
##      "< 0.05"      "< 0.05"      "< 0.05"      "< 0.05"      "< 0.05"      "< 0.05"
##      sp_attack      sp_defense      speed
##      "< 0.05"      "< 0.05"      "< 0.05"
```

Un cop vistos els resultats, es comprova que no hi ha cap variable que segueix una distribució normal, s'aplicarà el test de Spearman per calcular les correlacions entre parells de variables:

```
corr.general<-cor(pokedexAnalysis[,analisi_normalitat], method="spearman")
corr.general
```

```
##      height_m      weight_kg total_points      hp      attack      defense
## height_m      1.0000000  0.8445423    0.7271834  0.6421040  0.6229499  0.50081658
## weight_kg      0.8445423  1.0000000    0.6429082  0.6107825  0.5870074  0.53268085
## total_points  0.7271834  0.6429082    1.0000000  0.7224800  0.7253079  0.67868217
## hp            0.6421040  0.6107825    0.7224800  1.0000000  0.5793900  0.45752740
## attack        0.6229499  0.5870074    0.7253079  0.5793900  1.0000000  0.51463234
## defense       0.5008166  0.5326808    0.6786822  0.4575274  0.5146323  1.00000000
## sp_attack     0.4750580  0.3411488    0.7293632  0.4737787  0.3535374  0.31669441
## sp_defense    0.4941680  0.4432568    0.7464024  0.4922686  0.3216803  0.59515716
## speed         0.3278843  0.1916601    0.5454881  0.2674212  0.3658956  0.06822903
##      sp_attack sp_defense      speed
## height_m      0.4750580  0.4941680  0.32788433
## weight_kg      0.3411488  0.4432568  0.19166008
## total_points  0.7293632  0.7464024  0.54548811
## hp            0.4737787  0.4922686  0.26742123
## attack        0.3535374  0.3216803  0.36589561
## defense       0.3166944  0.5951572  0.06822903
## sp_attack     1.0000000  0.5783745  0.43790817
## sp_defense    0.5783745  1.0000000  0.27996316
## speed         0.4379082  0.2799632  1.00000000
```

No s'identifica cap correlació negativa. El parell de variables que estan més correlacionades són weight\_kg i height\_m. La que menys defense i speed.

```
apply(pokedexAnalysis.Legendary[,analisi_normalitat], 2, shapirotest)
```

### Compravació de la normalitat dels Llegendaris

```
##      height_m      weight_kg total_points      hp      attack      defense
##      "< 0.05"      "< 0.05"      "< 0.05"      "< 0.05"      "< 0.05"      "< 0.05"
##      sp_attack      sp_defense      speed
##      "> 0.05"      "< 0.05"      "< 0.05"
```

Un cop vistos els resultats, es comprova que hi ha una variable que compleix la distribució normal: sp\_attack. La resta no segueixen una distribució normal, s'aplicarà el test de Spearman per calcular les correlacions entre parells de variables:

```
corr.legendary<-cor(pokedexAnalysis.Legendary[,analisi_normalitat], method="spearman")
corr.legendary
```

```
##      height_m      weight_kg total_points      hp      attack
## height_m      1.00000000  0.397983487  0.39192759  0.45774863  0.053911459
## weight_kg      0.39798349  1.000000000  0.21754689  0.04250495  0.002593459
## total_points  0.39192759  0.217546890  1.00000000  0.33977698  0.539091466
## hp            0.45774863  0.042504951  0.33977698  1.00000000 -0.043127258
## attack        0.05391146  0.002593459  0.53909147 -0.04312726  1.000000000
## defense       0.13847321  0.622322099  0.08670783 -0.07119840  0.046906159
## sp_attack     0.37588954 -0.021325265  0.53442491  0.22852115  0.348675229
## sp_defense    0.17587644  0.217866441  0.27255534 -0.08076060 -0.068372144
## speed        -0.04348917 -0.366430094  0.37948600  0.09000829  0.264254457
##      defense      sp_attack      sp_defense      speed
## height_m      0.13847321  0.37588954  0.17587644 -0.04348917
## weight_kg      0.62232210 -0.02132527  0.21786644 -0.36643009
## total_points  0.08670783  0.53442491  0.27255534  0.37948600
## hp           -0.07119840  0.22852115 -0.08076060  0.09000829
## attack        0.04690616  0.34867523 -0.06837214  0.26425446
## defense       1.00000000 -0.38957318  0.22458477 -0.13148454
## sp_attack     -0.38957318  1.00000000  0.06694209  0.09998229
## sp_defense    0.22458477  0.06694209  1.00000000  0.07910543
## speed        -0.13148454  0.09998229  0.07910543  1.00000000
```

Hi ha bastantes correlacions negatives poc significatives. El parell de variables que estan més correlacionades són el sp\_attack i total\_points. La que menys sp\_defense i sp\_attack.

```
apply(pokedexAnalysis.SubLegendary[,analisi_normalitat], 2, shapirotest)
```

### Compravació de la normalitat dels Sub Llegendaris

```
##      height_m      weight_kg total_points      hp      attack      defense
##      "< 0.05"      "< 0.05"      "< 0.05"      "< 0.05"      "> 0.05"      "< 0.05"
##      sp_attack      sp_defense      speed
##      "> 0.05"      "> 0.05"      "< 0.05"
```

Un cop vistos els resultats, es comprova que hi ha variables que compleixen la distribució normal: attack, sp\_attack, sp\_defense, la resta no segueixen una distribució normal, s'aplicarà el test de Spearman per calcular les correlacions entre parells de variables:

```
corr.sblegendary<-cor(pokedexAnalysis.SubLegendary[,analisi_normalitat], method="spearman")
corr.sblegendary
```

```
##           height_m  weight_kg total_points           hp      attack
## height_m      1.00000000  0.71482100 -0.10937888  0.3871261  0.16353412
## weight_kg      0.71482100  1.00000000   0.11100220  0.6388801  0.03552098
## total_points -0.10937888  0.11100220   1.00000000  0.2828191  0.05984086
## hp            0.38712613  0.63888006   0.28281908  1.0000000 -0.13130977
## attack        0.16353412  0.03552098   0.05984086 -0.1313098  1.00000000
## defense       0.06658291  0.25794445   0.22539803  0.1094882  0.02799723
## sp_attack      0.00493100 -0.27465462   0.14757320 -0.1512102 -0.04418428
## sp_defense    -0.13084183  0.05104435   0.42070355  0.1621842 -0.51125440
## speed        -0.18535069 -0.36658769   0.21668816 -0.2527418  0.18146071
##           defense  sp_attack  sp_defense      speed
## height_m      0.06658291  0.00493100 -0.13084183 -0.1853507
## weight_kg      0.25794445 -0.27465462  0.05104435 -0.3665877
## total_points  0.22539803  0.14757320  0.42070355  0.2166882
## hp            0.10948823 -0.15121015  0.16218425 -0.2527418
## attack        0.02799723 -0.04418428 -0.51125440  0.1814607
## defense       1.00000000 -0.60343810  0.39018105 -0.4554884
## sp_attack     -0.60343810  1.00000000 -0.03616290  0.4414847
## sp_defense    0.39018105 -0.03616290  1.00000000 -0.2186953
## speed        -0.45548840  0.44148474 -0.21869530  1.0000000
```

Es pot observar que hi ha bastantes correlacions negatives poc significatives. El parell de variables que estan més correlacionades són el weight\_kg i el height\_m. La que menys sp\_attack i height\_m.

```
apply(pokedexAnalysis.Normal[,analisi_normalitat], 2, shapirotest)
```

### Compravació de la normalitat dels Normal

```
##      height_m  weight_kg total_points           hp      attack      defense
##      "< 0.05"    "< 0.05"    "< 0.05"    "< 0.05"    "< 0.05"    "< 0.05"
##      sp_attack  sp_defense      speed
##      "< 0.05"    "< 0.05"    "< 0.05"
```

Un cop vistos els resultats, es comprova que no hi ha cap variable que segueix una distribució normal, s'aplicarà el test de Spearman per calcular les correlacions entre parells de variables:

```
corr.normal<-cor(pokedexAnalysis.Normal[,analisi_normalitat], method="spearman")
corr.normal
```

```
##           height_m  weight_kg total_points           hp      attack      defense
## height_m      1.0000000  0.8247715   0.7294372  0.6406068  0.6087376  0.49122801
## weight_kg      0.8247715  1.0000000   0.6430732  0.6048420  0.5854619  0.51876330
```

```
## total_points 0.7294372 0.6430732 1.0000000 0.7025326 0.6892365 0.66387254
## hp 0.6406068 0.6048420 0.7025326 1.0000000 0.5569196 0.42615584
## attack 0.6087376 0.5854619 0.6892365 0.5569196 1.0000000 0.49535542
## defense 0.4912280 0.5187633 0.6638725 0.4261558 0.4953554 1.00000000
## sp_attack 0.4257457 0.2898501 0.6768441 0.4164971 0.2645908 0.27263737
## sp_defense 0.4787056 0.4163457 0.7299270 0.4515667 0.2715467 0.56427825
## speed 0.2793456 0.1440665 0.4658765 0.1905804 0.2876544 -0.01377796
## sp_attack sp_defense speed
## height_m 0.4257457 0.4787056 0.27934563
## weight_kg 0.2898501 0.4163457 0.14406653
## total_points 0.6768441 0.7299270 0.46587645
## hp 0.4164971 0.4515667 0.19058041
## attack 0.2645908 0.2715467 0.28765440
## defense 0.2726374 0.5642783 -0.01377796
## sp_attack 1.0000000 0.5467666 0.35072172
## sp_defense 0.5467666 1.0000000 0.20985845
## speed 0.3507217 0.2098584 1.00000000
```

Es pot observar que hi ha poques correlacions negatives poc significatives. El parell de variables que estan més correlacionades són el weight\_kg i el height\_m. La que menys sp\_defense i speed.

```
apply(pokedexAnalysis.Bug[,analisi_normalitat], 2, shapirotest)
```

### Comprovació de la normalitat dels Bug

```
## height_m weight_kg total_points hp attack defense
## "< 0.05" "< 0.05" "< 0.05" "> 0.05" "< 0.05" "< 0.05"
## sp_attack sp_defense speed
## "< 0.05" "< 0.05" "< 0.05"
```

Un cop vistos els resultats, es comprova que hi ha una variable que compleix la distribució normal: hp. La resta no segueixen una distribució normal, s'aplicarà el test de Spearman per calcular les correlacions entre parells de variables:

```
corr.bug<-cor(pokedexAnalysis.Bug[,analisi_normalitat], method="spearman")
corr.bug
```

```
## height_m weight_kg total_points hp attack defense
## height_m 1.0000000 0.8437690 0.7040183 0.7377776 0.6478494 0.38552800
## weight_kg 0.8437690 1.0000000 0.7367446 0.6937745 0.6279671 0.55983381
## total_points 0.7040183 0.7367446 1.0000000 0.8134250 0.7478347 0.63848409
## hp 0.7377776 0.6937745 0.8134250 1.0000000 0.7039882 0.42063760
## attack 0.6478494 0.6279671 0.7478347 0.7039882 1.0000000 0.47323396
## defense 0.3855280 0.5598338 0.6384841 0.4206376 0.4732340 1.00000000
## sp_attack 0.4660880 0.3975147 0.6814673 0.7223940 0.4660845 0.31791174
## sp_defense 0.4737494 0.5357487 0.7173375 0.5109216 0.3938180 0.58026635
## speed 0.4147443 0.3741652 0.5637374 0.5043968 0.4089315 -0.05941105
## sp_attack sp_defense speed
## height_m 0.4660880 0.4737494 0.41474434
## weight_kg 0.3975147 0.5357487 0.37416515
```

```
## total_points 0.6814673 0.7173375 0.56373745
## hp           0.7223940 0.5109216 0.50439681
## attack       0.4660845 0.3938180 0.40893152
## defense      0.3179117 0.5802663 -0.05941105
## sp_attack    1.0000000 0.5015327 0.48380282
## sp_defense   0.5015327 1.0000000 0.23134780
## speed        0.4838028 0.2313478 1.00000000
```

Es pot observar que hi ha poques correlacions negatives poc significatives. El parell de variables que estan més correlacionades són el weight\_kg i el height\_m. La que menys sp\_defense i speed.

```
apply(pokedexAnalysis.Dragon[,analisi_normalitat], 2, shapirotest)
```

### Compravació de la normalitat dels Dragon

```
##      height_m      weight_kg total_points      hp      attack      defense
##      "< 0.05"      "< 0.05"      "< 0.05"      "< 0.05"      "> 0.05"      "> 0.05"
##      sp_attack      sp_defense      speed
##      "> 0.05"      "> 0.05"      "> 0.05"
```

Un cop vistos els resultats, es comprova que hi ha variables que compleixen la distribució normal: attack, defense, sp\_attack, sp\_defense, speed, i la resta no segueixen una distribució normal, s'aplicarà el test de Spearman per calcular les correlacions entre parells de variables:

```
corr.dragon<-cor(pokedexAnalysis.Dragon[,analisi_normalitat], method="spearman")
corr.dragon
```

```
##      height_m weight_kg total_points      hp      attack      defense
## height_m    1.0000000 0.7347459 0.7893448 0.8095194 0.6738980 0.5690797
## weight_kg    0.7347459 1.0000000 0.7822978 0.8438520 0.7635649 0.7108920
## total_points 0.7893448 0.7822978 1.0000000 0.8965642 0.7753509 0.7963697
## hp           0.8095194 0.8438520 0.8965642 1.0000000 0.7728194 0.6904454
## attack       0.6738980 0.7635649 0.7753509 0.7728194 1.0000000 0.6523358
## defense      0.5690797 0.7108920 0.7963697 0.6904454 0.6523358 1.0000000
## sp_attack    0.7426935 0.6413536 0.9075369 0.7789917 0.6584228 0.6404532
## sp_defense   0.4822250 0.4338082 0.7130383 0.5745447 0.3513998 0.5814557
## speed        0.5256522 0.3675060 0.6775899 0.5707225 0.5638574 0.4351033
##      sp_attack sp_defense      speed
## height_m    0.7426935 0.4822250 0.5256522
## weight_kg    0.6413536 0.4338082 0.3675060
## total_points 0.9075369 0.7130383 0.6775899
## hp           0.7789917 0.5745447 0.5707225
## attack       0.6584228 0.3513998 0.5638574
## defense      0.6404532 0.5814557 0.4351033
## sp_attack    1.0000000 0.7877817 0.6436413
## sp_defense   0.7877817 1.0000000 0.4090417
## speed        0.6436413 0.4090417 1.0000000
```

Es pot observar que no hi ha correlacions negatives. El parell de variables que estan més correlacionades són el sp\_attack i total\_points. La que menys attack i sp\_defense.

```
apply(pokedexAnalysis.Steel[,analisi_normalitat], 2, shapirotest)
```

### Compravació de la normalitat dels Steel

```
##      height_m      weight_kg total_points      hp      attack      defense
##      "< 0.05"      "< 0.05"      "> 0.05"      "< 0.05"      "> 0.05"      "< 0.05"
##      sp_attack      sp_defense      speed
##      "< 0.05"      "> 0.05"      "< 0.05"
```

Un cop vistos els resultats, es comprova que hi ha variables que compleixen la distribució normal: total\_points, attack, sp\_defense, i la resta no segueixen una distribució normal, s'aplicarà el test de Spearman per calcular les correlacions entre parells de variables:

```
corr.steel<-cor(pokedexAnalysis.Steel[,analisi_normalitat], method="spearman")
corr.steel
```

```
##      height_m      weight_kg total_points      hp      attack      defense
## height_m      1.0000000 0.81106274      0.6974041 0.7199121 0.5801847 0.6091874
## weight_kg      0.8110627 1.00000000      0.5855208 0.6721461 0.4558946 0.4687492
## total_points 0.6974041 0.58552084      1.0000000 0.8046392 0.7245656 0.6544974
## hp      0.7199121 0.67214612      0.8046392 1.0000000 0.6128257 0.4151613
## attack      0.5801847 0.45589459      0.7245656 0.6128257 1.0000000 0.3928133
## defense      0.6091874 0.46874918      0.6544974 0.4151613 0.3928133 1.0000000
## sp_attack      0.4393772 0.38285900      0.7770171 0.6744308 0.6274037 0.1680696
## sp_defense      0.3503098 0.40413777      0.6309232 0.4926555 0.1683993 0.4182278
## speed      0.1598787 0.05476435      0.5891575 0.4181469 0.3519767 0.1930840
##      sp_attack sp_defense      speed
## height_m      0.4393772 0.3503098 0.15987866
## weight_kg      0.3828590 0.4041378 0.05476435
## total_points 0.7770171 0.6309232 0.58915747
## hp      0.6744308 0.4926555 0.41814689
## attack      0.6274037 0.1683993 0.35197669
## defense      0.1680696 0.4182278 0.19308396
## sp_attack      1.0000000 0.4832484 0.66285267
## sp_defense      0.4832484 1.0000000 0.42909192
## speed      0.6628527 0.4290919 1.00000000
```

Es pot observar que no hi ha correlacions negatives. El parell de variables que estan més correlacionades són el weight\_kg i height\_m. La que menys speed i weight\_kg.

```
apply(pokedexAnalysis.Poison[,analisi_normalitat], 2, shapirotest)
```

### Compravació de la normalitat dels Poison

```
##      height_m      weight_kg total_points      hp      attack      defense
##      "< 0.05"      "< 0.05"      "< 0.05"      "< 0.05"      "> 0.05"      "< 0.05"
##      sp_attack      sp_defense      speed
##      "< 0.05"      "< 0.05"      "< 0.05"
```

Un cop vistos els resultats, es comprova que hi ha una variable que compleix la distribució normal: attack. La resta no segueixen una distribució normal, s'aplicarà el test de Spearman per calcular les correlacions entre parells de variables:

```
corr.poison<-cor(pokedexAnalysis.Poison[,analisi_normalitat], method="spearman")
corr.poison
```

```
##           height_m weight_kg total_points      hp      attack      defense
## height_m      1.0000000 0.7402520    0.6928509 0.5095222 0.6855361 0.5120296
## weight_kg      0.7402520 1.0000000    0.7136130 0.7318304 0.6686769 0.4432401
## total_points  0.6928509 0.7136130    1.0000000 0.7142207 0.7621680 0.7857976
## hp            0.5095222 0.7318304    0.7142207 1.0000000 0.7628882 0.3893878
## attack        0.6855361 0.6686769    0.7621680 0.7628882 1.0000000 0.5254831
## defense       0.5120296 0.4432401    0.7857976 0.3893878 0.5254831 1.0000000
## sp_attack     0.6484030 0.5314066    0.7509551 0.5605338 0.6100252 0.4839936
## sp_defense    0.6637541 0.6925583    0.8661734 0.6330220 0.6815553 0.7906838
## speed        0.5784235 0.5302384    0.5798126 0.3975836 0.4134484 0.2834602
##           sp_attack sp_defense      speed
## height_m      0.6484030 0.6637541 0.5784235
## weight_kg      0.5314066 0.6925583 0.5302384
## total_points  0.7509551 0.8661734 0.5798126
## hp            0.5605338 0.6330220 0.3975836
## attack        0.6100252 0.6815553 0.4134484
## defense       0.4839936 0.7906838 0.2834602
## sp_attack     1.0000000 0.5841011 0.5776554
## sp_defense    0.5841011 1.0000000 0.3752602
## speed        0.5776554 0.3752602 1.0000000
```

Es pot observar que no hi ha correlacions negatives. El parell de variables que estan més correlacionades són el total\_points i sp\_defense. La que menys speed i defense.

```
apply(pokedexAnalysis.Psychic[,analisi_normalitat], 2, shapirotest)
```

### Comprovació de la normalitat dels Psychic

```
##      height_m      weight_kg total_points      hp      attack      defense
##      "< 0.05"      "< 0.05"      "> 0.05"      "< 0.05"      "< 0.05"      "< 0.05"
##      sp_attack      sp_defense      speed
##      "> 0.05"      "> 0.05"      "> 0.05"
```

Un cop vistos els resultats, es comprova que hi ha variables que compleixen la distribució normal: total\_points, sp\_attack, sp\_defense, speed, i la resta no segueixen una distribució normal, s'aplicarà el test de Spearman per calcular les correlacions entre parells de variables:

```
corr.psychic<-cor(pokedexAnalysis.Psychic[,analisi_normalitat], method="spearman")
corr.psychic
```

```
##           height_m weight_kg total_points      hp      attack      defense
## height_m      1.0000000 0.8287837    0.6798441 0.3613326 0.5936483 0.3732094
```



```
## weight_kg      0.8287837 1.0000000      0.5575698 0.2950586 0.4152088 0.3430801
## total_points  0.6798441 0.5575698      1.0000000 0.6304778 0.8883465 0.7051373
## hp            0.3613326 0.2950586      0.6304778 1.0000000 0.5269098 0.6331338
## attack        0.5936483 0.4152088      0.8883465 0.5269098 1.0000000 0.5751413
## defense       0.3732094 0.3430801      0.7051373 0.6331338 0.5751413 1.0000000
## sp_attack     0.5229052 0.3479175      0.7025825 0.3525706 0.6004285 0.2706368
## sp_defense    0.4430197 0.4860795      0.6761585 0.4349801 0.4578954 0.6822322
## speed        0.4282293 0.3190631      0.6543700 0.1634030 0.5790353 0.2810258
##              sp_attack sp_defense      speed
## height_m     0.5229052 0.4430197 0.4282293
## weight_kg    0.3479175 0.4860795 0.3190631
## total_points 0.7025825 0.6761585 0.6543700
## hp          0.3525706 0.4349801 0.1634030
## attack       0.6004285 0.4578954 0.5790353
## defense      0.2706368 0.6822322 0.2810258
## sp_attack    1.0000000 0.3535881 0.4735736
## sp_defense   0.3535881 1.0000000 0.2653745
## speed       0.4735736 0.2653745 1.0000000
```

Es pot observar que no hi ha correlacions negatives. El parell de variables que estan més correlacionades són el `weight_kg` i `height_m`. La que menys `speed` i `hp`.

**Comprovació de la homogenietat** Per mirar la homogenietat, utilitzarem el test *Fligner-killen* per comparar les variables amb distribució normal, sinó utilitzarem el test *Kruskal-wallis*.

```
## definim taula de variables per els càlculs de la normalitat i de la homogeneïtat de la variància
analisi_homogenietat <- c("height_m", "weight_kg", "hp", "attack", "defense", "sp_attack", "sp_defense",
total_points <- c("height_m", "weight_kg", "hp", "attack", "defense", "sp_attack", "sp_defense", "speed")

## funció per calcular el test de kruskal-wallis
kruskaltest <- function(y,x){
  alpha = 0.05
  p_val = kruskal.test(x ~ y)$p.value
  if (p_val > alpha) { "> 0.05"}
  else { "< 0.05"}
}
```

**Comprovació de la homogenietat de tot el grup** Com que totes les variables no tenien una distribució normal, aplicarem el test de Kruskal-wallis per veure el seu comportament, entre el `total_points` i la resta de variables:

```
sapply(pokedexAnalysis[,analisi_normalitat], kruskaltest, pokedexAnalysis$total_points)
```

```
##      height_m      weight_kg total_points      hp      attack      defense
##      "< 0.05"      "< 0.05"      "< 0.05"      "< 0.05"      "< 0.05"      "< 0.05"
##      sp_attack  sp_defense      speed
##      "< 0.05"      "< 0.05"      "< 0.05"
```

Es pot observar que cap variable la  $p > 0.05$ . No podem acceptar la hipòtesis de que les variances del total de punts amb la resta són homogènies.

**Comprovació de la homogenietat de Llegendaris** Com que totes les variables no tenien una distribució normal, aplicarem el test de Kruskal-wallis per veure el seu comportament, entre el total\_points i la resta de variables:

```
sapply(pokedexAnalysis.Legendary[,analisi_homogenietat], kruskaltest, pokedexAnalysis.Legendary$total_p
```

```
## height_m weight_kg hp attack defense sp_attack sp_defense
## "> 0.05" "> 0.05" "> 0.05" "> 0.05" "> 0.05" "> 0.05" "> 0.05"
## speed
## "> 0.05"
```

Es pot observar que totes les variables la  $p > 0.05$ . Podem acceptar la hipòtesis de que les variances del total de punts amb la resta són homogènies.

**Comprovació de la homogenietat de Sub Llegendaris** Com que totes les variables no tenien una distribució normal, aplicarem el test de Kruskal-wallis per veure el seu comportament, entre el total\_points i la resta de variables:

```
sapply(pokedexAnalysis.SubLegendary[,analisi_homogenietat], kruskaltest, pokedexAnalysis.SubLegendary$total_p
```

```
## height_m weight_kg hp attack defense sp_attack sp_defense
## "> 0.05" "> 0.05" "< 0.05" "> 0.05" "> 0.05" "> 0.05" "> 0.05"
## speed
## "> 0.05"
```

Es pot observar que totes les variables menys la “hp”, la  $p > 0.05$ . Podem acceptar la hipòtesis de que les variances del total de punts amb la resta són homogènies.

**Comprovació de la homogenietat de Normal** Com que totes les variables no tenien una distribució normal, aplicarem el test de Kruskal-wallis per veure el seu comportament, entre el total\_points i la resta de variables:

```
sapply(pokedexAnalysis.Normal[,analisi_homogenietat], kruskaltest, pokedexAnalysis.Normal$total_points)
```

```
## height_m weight_kg hp attack defense sp_attack sp_defense
## "< 0.05" "< 0.05" "< 0.05" "< 0.05" "< 0.05" "< 0.05" "< 0.05"
## speed
## "< 0.05"
```

Es pot observar que cap variable la  $p > 0.05$ . No podem acceptar la hipòtesis de que les variances del total de punts amb la resta són homogènies.

**Comprovació de la homogenietat del tipus Bug** Com que totes les variables no tenien una distribució normal, aplicarem el test de Kruskal-wallis per veure el seu comportament, entre el total\_points i la resta de variables:

```
sapply(pokedexAnalysis.Bug[,analisi_homogenietat], kruskaltest, pokedexAnalysis.Bug$total_points)
```

```
## height_m weight_kg hp attack defense sp_attack sp_defense
## "< 0.05" "> 0.05" "< 0.05" "< 0.05" "< 0.05" "< 0.05" "< 0.05"
## speed
## "< 0.05"
```

Es pot observar que només hi ha la variable “weight\_kg” on la  $p > 0.05$ . Només podem acceptar la hipòtesis de que les variàncies del total de punts amb la resta són homogènies, en la variable “weight\_kg” respecte “total\_points”.

**Comprovació de la homogenietat del tipus Dragon** Com que totes les variables no tenien una distribució normal, aplicarem el test de Kruskal-wallis per veure el seu comportament, entre el total\_points i la resta de variables:

```
supply(pokedexAnalysis.Dragon[,analisi_homogenietat], kruskaltest, pokedexAnalysis.Dragon$total_points)
```

```
## height_m weight_kg hp attack defense sp_attack sp_defense
## "> 0.05" "> 0.05" "< 0.05" "> 0.05" "> 0.05" "< 0.05" "< 0.05"
## speed
## "> 0.05"
```

Es pot observar que totes les variables menys la “hp”, “sp\_attack” i “sp\_defense”, la  $p > 0.05$ . Podem acceptar la hipòtesis de que les variàncies del total de punts amb la resta són homogènies, en les variables que tenen la  $p > 0.05$ .

**Comprovació de la homogenietat del tipus Steel** Com que totes les variables no tenien una distribució normal, aplicarem el test de Kruskal-wallis per veure el seu comportament, entre el total\_points i la resta de variables:

```
supply(pokedexAnalysis.Steel[,analisi_homogenietat], kruskaltest, pokedexAnalysis.Steel$total_points)
```

```
## height_m weight_kg hp attack defense sp_attack sp_defense
## "> 0.05" "> 0.05" "< 0.05" "> 0.05" "> 0.05" "> 0.05" "> 0.05"
## speed
## "> 0.05"
```

Es pot observar que totes les variables menys la “hp”, la  $p > 0.05$ . Podem acceptar la hipòtesis de que les variàncies del total de punts amb la resta són homogènies.

**Comprovació de la homogenietat del tipus Poison** Com que totes les variables no tenien una distribució normal, aplicarem el test de Kruskal-wallis per veure el seu comportament, entre el total\_points i la resta de variables:

```
supply(pokedexAnalysis.Poison[,analisi_homogenietat], kruskaltest, pokedexAnalysis.Poison$total_points)
```

```
## height_m weight_kg hp attack defense sp_attack sp_defense
## "< 0.05" "> 0.05" "> 0.05" "> 0.05" "> 0.05" "< 0.05" "> 0.05"
## speed
## "> 0.05"
```

Es pot observar que totes les variables menys la “height\_m” i “sp\_attack”, la  $p > 0.05$ . Podem acceptar la hipòtesis de que les variàncies del total de punts amb la resta són homogènies, només en les que la  $p > 0.05$ .

**Comprovació de la homogenietat del tipus Psychic** Com que totes les variables no tenien una distribució normal, aplicarem el test de Kruskal-wallis per veure el seu comportament, entre el total\_points i la resta de variables:

```
sapply(pokedexAnalysis.Psychic[,analisi_homogenietat], kruskaltest, pokedexAnalysis.Psychic$total_points)
```

```
## height_m weight_kg hp attack defense sp_attack sp_defense
## "< 0.05" "> 0.05" "< 0.05" "< 0.05" "< 0.05" "> 0.05" "< 0.05"
## speed
## "< 0.05"
```

Es pot observar que només les variables “weight\_kg” i “sp\_attack”, la  $p > 0.05$ . Podem acceptar la hipòtesis de que les variàncies del total de punts amb la resta són homogènies, només en les que la  $p > 0.05$ .

**4.3. Aplicació de proves estadístiques per comparar els grups de dades. En funció de les dades i de l'objectiu de l'estudi, aplicar proves de contrast d'hipòtesis, correlacions, regressions, etc. Aplicar almenys tres mètodes d'anàlisi diferents.**

**4.3.1. Estudiar les variables d'estadística base, que fan determinar la força total d'un Pokemon. Es vol estudiar la relació entre aquests atributs. S'estudiarà de forma global entre tots els tipus i llavors amb la classe Dragon i Bug, que són els més forts i dèbils respectivament. També estudiarem les classes Steel, Poison, Psychic. Influx igual en els Pokemons llegendaris i sub llegendaris?** Com que per estudiar la normalitat de les dades s'han realitzat les matrius de correlació, les mostrarem i estudiarem la importància de les variables segons el total\_points:

```
corr.general
```

Matriu de correlació dels Pokemons en general

```
## height_m weight_kg total_points hp attack defense
## height_m 1.000000 0.8445423 0.7271834 0.6421040 0.6229499 0.50081658
## weight_kg 0.8445423 1.0000000 0.6429082 0.6107825 0.5870074 0.53268085
## total_points 0.7271834 0.6429082 1.0000000 0.7224800 0.7253079 0.67868217
## hp 0.6421040 0.6107825 0.7224800 1.0000000 0.5793900 0.45752740
## attack 0.6229499 0.5870074 0.7253079 0.5793900 1.0000000 0.51463234
## defense 0.5008166 0.5326808 0.6786822 0.4575274 0.5146323 1.00000000
## sp_attack 0.4750580 0.3411488 0.7293632 0.4737787 0.3535374 0.31669441
## sp_defense 0.4941680 0.4432568 0.7464024 0.4922686 0.3216803 0.59515716
## speed 0.3278843 0.1916601 0.5454881 0.2674212 0.3658956 0.06822903
## sp_attack sp_defense speed
## height_m 0.4750580 0.4941680 0.32788433
## weight_kg 0.3411488 0.4432568 0.19166008
## total_points 0.7293632 0.7464024 0.54548811
## hp 0.4737787 0.4922686 0.26742123
## attack 0.3535374 0.3216803 0.36589561
## defense 0.3166944 0.5951572 0.06822903
## sp_attack 1.0000000 0.5783745 0.43790817
## sp_defense 0.5783745 1.0000000 0.27996316
## speed 0.4379082 0.2799632 1.00000000
```

En base a la variable total\_points, es pot veure que amb la variable que està més correlacionada és la de sp\_defense, i llavors les quatre variables següents es podria dir que estan igual de correlacionades, sp\_attack, height\_m, attack i hp.

sp\_defense: 0.7464024 sp\_attack: 0.7293632 height\_m: 0.7271834 attack: 0.7253079 hp: 0.7224800 defense: 0.6786822 weight\_kg: 0.6429082 speed: 0.5454881

```
corr.normal
```

### Matriu de correlació dels Pokemons de l'estatus normal

```
##          height_m weight_kg total_points          hp    attack    defense
## height_m    1.0000000 0.8247715    0.7294372 0.6406068 0.6087376 0.49122801
## weight_kg    0.8247715 1.0000000    0.6430732 0.6048420 0.5854619 0.51876330
## total_points 0.7294372 0.6430732    1.0000000 0.7025326 0.6892365 0.66387254
## hp           0.6406068 0.6048420    0.7025326 1.0000000 0.5569196 0.42615584
## attack       0.6087376 0.5854619    0.6892365 0.5569196 1.0000000 0.49535542
## defense      0.4912280 0.5187633    0.6638725 0.4261558 0.4953554 1.00000000
## sp_attack    0.4257457 0.2898501    0.6768441 0.4164971 0.2645908 0.27263737
## sp_defense   0.4787056 0.4163457    0.7299270 0.4515667 0.2715467 0.56427825
## speed        0.2793456 0.1440665    0.4658765 0.1905804 0.2876544 -0.01377796
##          sp_attack sp_defense          speed
## height_m    0.4257457 0.4787056 0.27934563
## weight_kg    0.2898501 0.4163457 0.14406653
## total_points 0.6768441 0.7299270 0.46587645
## hp           0.4164971 0.4515667 0.19058041
## attack       0.2645908 0.2715467 0.28765440
## defense      0.2726374 0.5642783 -0.01377796
## sp_attack    1.0000000 0.5467666 0.35072172
## sp_defense   0.5467666 1.0000000 0.20985845
## speed        0.3507217 0.2098584 1.00000000
```

En base a la variable total\_points, es pot veure que amb la variable que està més correlacionada és la de sp\_defense juntament amb height\_m.

sp\_defense: 0.7299270 height\_m: 0.7294372 hp: 0.7025326 attack: 0.6892365 sp\_attack: 0.6768441 defense: 0.6638725 weight\_kg: 0.6430732 speed: 0.4658765

```
corr.legendary
```

### Matriu de correlació dels Pokemons de l'estatus llegendari

```
##          height_m    weight_kg total_points          hp    attack
## height_m    1.00000000 0.397983487    0.39192759 0.45774863 0.053911459
## weight_kg    0.39798349 1.000000000    0.21754689 0.04250495 0.002593459
## total_points 0.39192759 0.217546890    1.00000000 0.33977698 0.539091466
## hp           0.45774863 0.042504951    0.33977698 1.00000000 -0.043127258
## attack       0.05391146 0.002593459    0.53909147 -0.04312726 1.000000000
```

```
## defense      0.13847321  0.622322099  0.08670783 -0.07119840  0.046906159
## sp_attack    0.37588954 -0.021325265  0.53442491  0.22852115  0.348675229
## sp_defense   0.17587644  0.217866441  0.27255534 -0.08076060 -0.068372144
## speed       -0.04348917 -0.366430094  0.37948600  0.09000829  0.264254457
##              defense  sp_attack  sp_defense    speed
## height_m     0.13847321  0.37588954  0.17587644 -0.04348917
## weight_kg     0.62232210 -0.02132527  0.21786644 -0.36643009
## total_points  0.08670783  0.53442491  0.27255534  0.37948600
## hp           -0.07119840  0.22852115 -0.08076060  0.09000829
## attack        0.04690616  0.34867523 -0.06837214  0.26425446
## defense       1.00000000 -0.38957318  0.22458477 -0.13148454
## sp_attack     -0.38957318  1.00000000  0.06694209  0.09998229
## sp_defense    0.22458477  0.06694209  1.00000000  0.07910543
## speed        -0.13148454  0.09998229  0.07910543  1.00000000
```

En base a la variable total\_points, es pot veure que amb la variable que està més correlacionada és la de attack juntament amb sp\_attack.

attack: 0.53909147 sp\_attack: 0.53442491 height\_m: 0.39192759 speed: 0.37948600 hp: 0.33977698  
sp\_defense: 0.27255534 weight\_kg: 0.21754689 defense: 0.08670783

```
corr.sublegendary
```

### Matriu de correlació dels Pokemons de l'estatus sub llegendaris

```
##              height_m  weight_kg total_points    hp    attack
## height_m     1.00000000  0.71482100 -0.10937888  0.3871261  0.16353412
## weight_kg     0.71482100  1.00000000  0.11100220  0.6388801  0.03552098
## total_points -0.10937888  0.11100220  1.00000000  0.2828191  0.05984086
## hp           0.38712613  0.63888006  0.28281908  1.0000000 -0.13130977
## attack       0.16353412  0.03552098  0.05984086 -0.1313098  1.00000000
## defense      0.06658291  0.25794445  0.22539803  0.1094882  0.02799723
## sp_attack    0.00493100 -0.27465462  0.14757320 -0.1512102 -0.04418428
## sp_defense   -0.13084183  0.05104435  0.42070355  0.1621842 -0.51125440
## speed       -0.18535069 -0.36658769  0.21668816 -0.2527418  0.18146071
##              defense  sp_attack  sp_defense    speed
## height_m     0.06658291  0.00493100 -0.13084183 -0.1853507
## weight_kg     0.25794445 -0.27465462  0.05104435 -0.3665877
## total_points  0.22539803  0.14757320  0.42070355  0.2166882
## hp           0.10948823 -0.15121015  0.16218425 -0.2527418
## attack       0.02799723 -0.04418428 -0.51125440  0.1814607
## defense      1.00000000 -0.60343810  0.39018105 -0.4554884
## sp_attack    -0.60343810  1.00000000 -0.03616290  0.4414847
## sp_defense   0.39018105 -0.03616290  1.00000000 -0.2186953
## speed       -0.45548840  0.44148474 -0.21869530  1.0000000
```

En base a la variable total\_points, es pot veure que amb la variable que està més correlacionada és la de sp\_defense.

sp\_defense: 0.42070355 hp: 0.28281908 defense: 0.22539803 speed: 0.21668816 sp\_attack: 0.14757320  
weight\_kg: 0.11100220 attack: 0.05984086

```
corr.bug
```

### Matriu de correlació dels Pokemons del tipus 1 bug

```
##          height_m weight_kg total_points          hp    attack    defense
## height_m    1.0000000 0.8437690    0.7040183 0.7377776 0.6478494 0.38552800
## weight_kg    0.8437690 1.0000000    0.7367446 0.6937745 0.6279671 0.55983381
## total_points 0.7040183 0.7367446    1.0000000 0.8134250 0.7478347 0.63848409
## hp           0.7377776 0.6937745    0.8134250 1.0000000 0.7039882 0.42063760
## attack       0.6478494 0.6279671    0.7478347 0.7039882 1.0000000 0.47323396
## defense      0.3855280 0.5598338    0.6384841 0.4206376 0.4732340 1.00000000
## sp_attack     0.4660880 0.3975147    0.6814673 0.7223940 0.4660845 0.31791174
## sp_defense    0.4737494 0.5357487    0.7173375 0.5109216 0.3938180 0.58026635
## speed         0.4147443 0.3741652    0.5637374 0.5043968 0.4089315 -0.05941105
##          sp_attack sp_defense          speed
## height_m    0.4660880 0.4737494 0.41474434
## weight_kg    0.3975147 0.5357487 0.37416515
## total_points 0.6814673 0.7173375 0.56373745
## hp           0.7223940 0.5109216 0.50439681
## attack       0.4660845 0.3938180 0.40893152
## defense      0.3179117 0.5802663 -0.05941105
## sp_attack     1.0000000 0.5015327 0.48380282
## sp_defense    0.5015327 1.0000000 0.23134780
## speed         0.4838028 0.2313478 1.00000000
```

En base a la variable total\_points, es pot veure que amb la variable que està més correlacionada és la de hp.

hp: 0.8134250 attack: 0.7478347 weight\_kg: 0.7367446 sp\_defense: 0.7173375 height\_m: 0.7040183  
sp\_attack: 0.6814673 defense: 0.6384841 speed: 0.5637374

```
corr.dragon
```

### Matriu de correlació dels Pokemons del tipus 1 dragon

```
##          height_m weight_kg total_points          hp    attack    defense
## height_m    1.0000000 0.7347459    0.7893448 0.8095194 0.6738980 0.5690797
## weight_kg    0.7347459 1.0000000    0.7822978 0.8438520 0.7635649 0.7108920
## total_points 0.7893448 0.7822978    1.0000000 0.8965642 0.7753509 0.7963697
## hp           0.8095194 0.8438520    0.8965642 1.0000000 0.7728194 0.6904454
## attack       0.6738980 0.7635649    0.7753509 0.7728194 1.0000000 0.6523358
## defense      0.5690797 0.7108920    0.7963697 0.6904454 0.6523358 1.0000000
## sp_attack     0.7426935 0.6413536    0.9075369 0.7789917 0.6584228 0.6404532
## sp_defense    0.4822250 0.4338082    0.7130383 0.5745447 0.3513998 0.5814557
## speed         0.5256522 0.3675060    0.6775899 0.5707225 0.5638574 0.4351033
##          sp_attack sp_defense          speed
## height_m    0.7426935 0.4822250 0.5256522
## weight_kg    0.6413536 0.4338082 0.3675060
## total_points 0.9075369 0.7130383 0.6775899
## hp           0.7789917 0.5745447 0.5707225
```

```
## attack      0.6584228  0.3513998  0.5638574
## defense     0.6404532  0.5814557  0.4351033
## sp_attack   1.0000000  0.7877817  0.6436413
## sp_defense  0.7877817  1.0000000  0.4090417
## speed       0.6436413  0.4090417  1.0000000
```

En base a la variable total\_points, es pot veure que amb la variable que està més correlacionada és la de sp\_attack juntament amb hp.

```
sp_attack 0.9075369 hp 0.8965642 defense 0.7963697 height_m 0.7893448 weight_kg 0.7822978 attack
0.7753509 sp_defense 0.7130383 speed 0.6775899
```

```
corr.poison
```

### Matriu de correlació dels Pokemons del tipus 1 poison

```
##          height_m weight_kg total_points      hp    attack    defense
## height_m  1.0000000  0.7402520    0.6928509 0.5095222 0.6855361 0.5120296
## weight_kg  0.7402520  1.0000000    0.7136130 0.7318304 0.6686769 0.4432401
## total_points 0.6928509 0.7136130    1.0000000 0.7142207 0.7621680 0.7857976
## hp         0.5095222 0.7318304    0.7142207 1.0000000 0.7628882 0.3893878
## attack     0.6855361 0.6686769    0.7621680 0.7628882 1.0000000 0.5254831
## defense    0.5120296 0.4432401    0.7857976 0.3893878 0.5254831 1.0000000
## sp_attack   0.6484030 0.5314066    0.7509551 0.5605338 0.6100252 0.4839936
## sp_defense  0.6637541 0.6925583    0.8661734 0.6330220 0.6815553 0.7906838
## speed       0.5784235 0.5302384    0.5798126 0.3975836 0.4134484 0.2834602
##          sp_attack sp_defense      speed
## height_m  0.6484030 0.6637541 0.5784235
## weight_kg  0.5314066 0.6925583 0.5302384
## total_points 0.7509551 0.8661734 0.5798126
## hp         0.5605338 0.6330220 0.3975836
## attack     0.6100252 0.6815553 0.4134484
## defense    0.4839936 0.7906838 0.2834602
## sp_attack   1.0000000 0.5841011 0.5776554
## sp_defense  0.5841011 1.0000000 0.3752602
## speed       0.5776554 0.3752602 1.0000000
```

En base a la variable total\_points, es pot veure que amb la variable que està més correlacionada és la de sp\_defense.

```
sp_defense 0.8661734 defense 0.7857976 attack 0.7621680 sp_attack 0.7509551 hp 0.7142207 weight_kg
0.7136130 height_m 0.6928509 speed 0.5798126
```

```
corr.steel
```

### Matriu de correlació dels Pokemons del tipus 1 steel



```
##          height_m weight_kg total_points      hp    attack    defense
## height_m    1.0000000 0.81106274    0.6974041 0.7199121 0.5801847 0.6091874
## weight_kg    0.8110627 1.00000000    0.5855208 0.6721461 0.4558946 0.4687492
## total_points 0.6974041 0.58552084    1.0000000 0.8046392 0.7245656 0.6544974
## hp          0.7199121 0.67214612    0.8046392 1.0000000 0.6128257 0.4151613
## attack      0.5801847 0.45589459    0.7245656 0.6128257 1.0000000 0.3928133
## defense     0.6091874 0.46874918    0.6544974 0.4151613 0.3928133 1.0000000
## sp_attack    0.4393772 0.38285900    0.7770171 0.6744308 0.6274037 0.1680696
## sp_defense   0.3503098 0.40413777    0.6309232 0.4926555 0.1683993 0.4182278
## speed       0.1598787 0.05476435    0.5891575 0.4181469 0.3519767 0.1930840
##            sp_attack sp_defense      speed
## height_m    0.4393772 0.3503098 0.15987866
## weight_kg    0.3828590 0.4041378 0.05476435
## total_points 0.7770171 0.6309232 0.58915747
## hp          0.6744308 0.4926555 0.41814689
## attack      0.6274037 0.1683993 0.35197669
## defense     0.1680696 0.4182278 0.19308396
## sp_attack    1.0000000 0.4832484 0.66285267
## sp_defense   0.4832484 1.0000000 0.42909192
## speed       0.6628527 0.4290919 1.00000000
```

En base a la variable total\_points, es pot veure que amb la variable que està més correlacionada és la de hp.

hp 0.8046392 sp\_attack 0.7770171 attack 0.7245656 height\_m 0.6974041 defense 0.6544974 sp\_defense 0.6309232 speed 0.5891575 weight\_kg 0.5855208

```
corr.psychic
```

### Matriu de correlació dels Pokemons del tipus 1 psychic

```
##          height_m weight_kg total_points      hp    attack    defense
## height_m    1.0000000 0.8287837    0.6798441 0.3613326 0.5936483 0.3732094
## weight_kg    0.8287837 1.0000000    0.5575698 0.2950586 0.4152088 0.3430801
## total_points 0.6798441 0.5575698    1.0000000 0.6304778 0.8883465 0.7051373
## hp          0.3613326 0.2950586    0.6304778 1.0000000 0.5269098 0.6331338
## attack      0.5936483 0.4152088    0.8883465 0.5269098 1.0000000 0.5751413
## defense     0.3732094 0.3430801    0.7051373 0.6331338 0.5751413 1.0000000
## sp_attack    0.5229052 0.3479175    0.7025825 0.3525706 0.6004285 0.2706368
## sp_defense   0.4430197 0.4860795    0.6761585 0.4349801 0.4578954 0.6822322
## speed       0.4282293 0.3190631    0.6543700 0.1634030 0.5790353 0.2810258
##            sp_attack sp_defense      speed
## height_m    0.5229052 0.4430197 0.4282293
## weight_kg    0.3479175 0.4860795 0.3190631
## total_points 0.7025825 0.6761585 0.6543700
## hp          0.3525706 0.4349801 0.1634030
## attack      0.6004285 0.4578954 0.5790353
## defense     0.2706368 0.6822322 0.2810258
## sp_attack    1.0000000 0.3535881 0.4735736
## sp_defense   0.3535881 1.0000000 0.2653745
## speed       0.4735736 0.2653745 1.0000000
```

En base a la variable `total_points`, es pot veure que amb la variable que està més correlacionada és la de `attack`.

```
attack 0.8883465 defense 0.7051373 sp_attack 0.7025825 height_m 0.6798441 sp_defense 0.6761585 speed
0.6543700 hp 0.6304778 weight_kg 0.5575698
```

**4.3.2. Estudiar si segon diferents tipus de pes i d'alçada del Pokemon, influeixen a que el Pokemon sigui més fort, per tant tinguin un total de punts més elevat.** La segona prova estadística que es realitzarà serà el contrast d'hipotesis sobre dues mostres per determinar si el `total_points` del Pokémon és superior depenent de la seva alçada, i si el `total_points` del Pokémon és superior depenent del seu pes.

**Test de dues mostres: el total de punts (força) del Pokémon en funció del pes**

**Test de dues mostres de Tots els Pokemons Segons pes:**

$H_0: \mu_{lleuger} = \mu_{pesat}$   $H_1: \mu_{lleuger} < \mu_{pesat}$

Utilitzarem un test de dues mostres independents. Pel teorema del límit central, assumim normalitat.

```
var.test(pokedexAnalysis$total_points[pokedexAnalysis$Fweight=="lleuger"], pokedexAnalysis$total_points

##
## F test to compare two variances
##
## data:  pokedexAnalysis$total_points[pokedexAnalysis$Fweight == "lleuger"] and pokedexAnalysis$total_
## F = 1.1227, num df = 769, denom df = 257, p-value = 0.8659
## alternative hypothesis: true ratio of variances is less than 1
## 95 percent confidence interval:
##  0.000000 1.322509
## sample estimates:
## ratio of variances
##          1.122715
```

El resultat del test F és que podem assumir igualtat de variàncies. Per tant, apliquem test t de dues mostres independents per a la diferència de mitjanes, variàncies desconegudes i iguals. El test és unilateral.

```
t.test(total_points~Fweight_kg,pokedexAnalysis,var.equal=T, alternative="less")

##
## Two Sample t-test
##
## data:  total_points by Fweight_kg
## t = -17.871, df = 1026, p-value < 2.2e-16
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##      -Inf -124.066
## sample estimates:
## mean in group lleuger    mean in group pesat
##          403.2753          539.9302
```

El valor p és molt inferior als nivells 0.05 amb un 95% de nivell de confiança, per tant, podem rebutjar la hipòtesis nul·la del qual el total de punts del Pokemons, són igual amb el pes lleuger i pesat. Es pot dir que els Pokemons amb més pes tenen un `total_points` superior.

### Test de dues mostres de Pokemons Llegendaris Segons pes:

H0:  $\mu_{lleuger} = \mu_{pesat}$  H1:  $\mu_{lleuger} < \mu_{pesat}$

Utilitzarem un test de dues mostres independents. Pel teorema del límit central, assumim normalitat.

```
var.test(pokedexAnalysis.Legendary$total_points[pokedexAnalysis.Legendary$Fweight=="lleuger"], pokedexAnalysis.Legendary$total_points[pokedexAnalysis.Legendary$Fweight=="pesat"],
data = pokedexAnalysis.Legendary,
var.equal = F,
alternative = "less")

##
## F test to compare two variances
##
## data:  pokedexAnalysis.Legendary$total_points[pokedexAnalysis.Legendary$Fweight == "lleuger"] and pokedexAnalysis.Legendary$total_points[pokedexAnalysis.Legendary$Fweight == "pesat"]
## F = 8.9756, num df = 2, denom df = 35, p-value = 0.9993
## alternative hypothesis: true ratio of variances is less than 1
## 95 percent confidence interval:
##      0.0000 174.7287
## sample estimates:
## ratio of variances
##              8.975558
```

El resultat del test F és que podem assumir igualtat de variàncies. Per tant, apliquem test t de dues mostres independents per a la diferència de mitjanes, variàncies desconegudes i iguals. El test és unilateral.

```
t.test(total_points~Fweight_kg,pokedexAnalysis.Legendary,var.equal=T, alternative="less")

##
## Two Sample t-test
##
## data:  total_points by Fweight_kg
## t = -2.9718, df = 37, p-value = 0.00259
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##      -Inf -89.40306
## sample estimates:
## mean in group lleuger    mean in group pesat
##           488.6667           695.4722
```

El valor p és inferior als nivells 0.05 amb un 95% de nivell de confiança, per tant, podem rebutjar la hipòtesis nul·la del qual el total de punts del Pokemons, són igual amb el pes lleuger i pesat. Es pot dir que els Pokemons amb més pes tenen un total\_points superior.

### Test de dues mostres de Pokemons Sub Llegendaris Segons pes:

H0:  $\mu_{lleuger} = \mu_{pesat}$  H1:  $\mu_{lleuger} < \mu_{pesat}$

Utilitzarem un test de dues mostres independents. Pel teorema del límit central, assumim normalitat.

```
var.test(pokedexAnalysis.SubLegendary$total_points[pokedexAnalysis.SubLegendary$Fweight=="lleuger"], pokedexAnalysis.SubLegendary$total_points[pokedexAnalysis.SubLegendary$Fweight=="pesat"],
data = pokedexAnalysis.SubLegendary,
var.equal = F,
alternative = "less")

##
## F test to compare two variances
##
## data:  pokedexAnalysis.SubLegendary$total_points[pokedexAnalysis.SubLegendary$Fweight == "lleuger"] and pokedexAnalysis.SubLegendary$total_points[pokedexAnalysis.SubLegendary$Fweight == "pesat"]
```

```
## F = 1.3583, num df = 23, denom df = 20, p-value = 0.7537
## alternative hypothesis: true ratio of variances is less than 1
## 95 percent confidence interval:
##  0.000000 2.781335
## sample estimates:
## ratio of variances
##           1.358314
```

El resultat del test F és que podem assumir igualtat de variàncies. Per tant, apliquem test t de dues mostres independents per a la diferència de mitjanes, variàncies desconegudes i iguals. El test és unilateral.

```
t.test(total_points~Fweight_kg,pokedexAnalysis.SubLegendary,var.equal=T, alternative="less")
```

```
##
## Two Sample t-test
##
## data: total_points by Fweight_kg
## t = -0.55125, df = 43, p-value = 0.2922
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
## -Inf 13.566
## sample estimates:
## mean in group lleuger mean in group pesat
##           578.3333           584.9524
```

El valor p és superior als nivells 0.05 amb un 95% de nivell de confiança, per tant, podem acceptar la hipòtesis nul·la del qual el total de punts del Pokemons, són igual amb el pes lleuger i pesat.

### Test de dues mostres de Pokemons de tipus Normal Segons pes:

H0:  $\mu_{lleuger} = \mu_{pesat}$  H1:  $\mu_{lleuger} < \mu_{pesat}$

Utilitzarem un test de dues mostres independents. Pel teorema del límit central, assumim normalitat.

```
var.test(pokedexAnalysis.Normal$total_points[pokedexAnalysis.Normal$Fweight=="lleuger"], pokedexAnalysis.No
```

```
##
## F test to compare two variances
##
## data: pokedexAnalysis.Normal$total_points[pokedexAnalysis.Normal$Fweight == "lleuger"] and pokedexAnalysis.No
## F = 1.7726, num df = 719, denom df = 194, p-value = 1
## alternative hypothesis: true ratio of variances is less than 1
## 95 percent confidence interval:
##  0.000000 2.127808
## sample estimates:
## ratio of variances
##           1.772595
```

El resultat del test F és que podem assumir igualtat de variàncies. Per tant, apliquem test t de dues mostres independents per a la diferència de mitjanes, variàncies desconegudes i iguals. El test és unilateral.

```
t.test(total_points~Fweight_kg,pokedexAnalysis.Normal,var.equal=T, alternative="less")
```

```
##
## Two Sample t-test
##
## data: total_points by Fweight_kg
## t = -14.787, df = 913, p-value < 2.2e-16
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##      -Inf -99.69994
## sample estimates:
## mean in group lleuger    mean in group pesat
##           391.3000           503.4923
```

El valor p és molt inferior als nivells 0.05 amb un 95% de nivell de confiança, per tant, podem rebutjar la hipòtesis nul·la del qual el total de punts dels Pokemons, són igual amb el pes lleuger i pesat. Es pot dir que els Pokemons amb més pes tenen un total\_points superior.

### Test de dues mostres de Pokemons de tipus Bug Segons pes:

$H_0: \mu_{lleuger} = \mu_{pesat}$   $H_1: \mu_{lleuger} < \mu_{pesat}$

Utilitzarem un test de dues mostres independents. Pel teorema del límit central, assumim normalitat.

```
var.test(pokedexAnalysis.Bug$total_points[pokedexAnalysis.Bug$Fweight=="lleuger"], pokedexAnalysis.Bug$
```

```
##
## F test to compare two variances
##
## data: pokedexAnalysis.Bug$total_points[pokedexAnalysis.Bug$Fweight == "lleuger"] and pokedexAnalysis.Bug$total_points[pokedexAnalysis.Bug$Fweight == "pesat"]
## F = 4.5195, num df = 72, denom df = 7, p-value = 0.9792
## alternative hypothesis: true ratio of variances is less than 1
## 95 percent confidence interval:
##  0.000000 9.670107
## sample estimates:
## ratio of variances
##           4.519469
```

El resultat del test F és que podem assumir igualtat de variàncies. Per tant, apliquem test t de dues mostres independents per a la diferència de mitjanes, variàncies desconegudes i iguals. El test és unilateral.

```
t.test(total_points~Fweight_kg,pokedexAnalysis.Bug,var.equal=T, alternative="less")
```

```
##
## Two Sample t-test
##
## data: total_points by Fweight_kg
## t = -3.8785, df = 79, p-value = 0.0001081
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##      -Inf -91.6941
## sample estimates:
## mean in group lleuger    mean in group pesat
##           368.7534           529.3750
```

El valor p és molt inferior als nivells 0.05 amb un 95% de nivell de confiança, per tant, podem rebutjar la hipòtesis nul·la del qual el total de punts del Pokemons, són igual amb el pes lleuger i pesat. Es pot dir que els Pokemons amb més pes tenen un total\_points superior.

### Test de dues mostres de Pokemons de tipus Dragon Segons pes:

H0:  $\mu_{\text{lleuger}} = \mu_{\text{pesat}}$  H1:  $\mu_{\text{lleuger}} < \mu_{\text{pesat}}$

Utilitzarem un test de dues mostres independents. Pel teorema del límit central, assumim normalitat.

```
var.test(pokedexAnalysis.Dragon$total_points[pokedexAnalysis.Dragon$Fweight=="lleuger"], pokedexAnalysis.Dragon$total_points[pokedexAnalysis.Dragon$Fweight=="pesat"],
data = pokedexAnalysis.Dragon,
var.equal=T, alternative="less")

##
## F test to compare two variances
##
## data:  pokedexAnalysis.Dragon$total_points[pokedexAnalysis.Dragon$Fweight == "lleuger"] and pokedexAnalysis.Dragon$total_points[pokedexAnalysis.Dragon$Fweight == "pesat"]
## F = 2.2677, num df = 19, denom df = 19, p-value = 0.9589
## alternative hypothesis: true ratio of variances is less than 1
## 95 percent confidence interval:
##  0.000000 4.916886
## sample estimates:
## ratio of variances
##          2.267673
```

El resultat del test F és que podem assumir igualtat de variàncies. Per tant, apliquem test t de dues mostres independents per a la diferència de mitjanes, variàncies desconegudes i iguals. El test és unilateral.

```
t.test(total_points~Fweight_kg,pokedexAnalysis.Dragon,var.equal=T, alternative="less")

##
## Two Sample t-test
##
## data:  total_points by Fweight_kg
## t = -5.8597, df = 38, p-value = 4.434e-07
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##      -Inf -144.4147
## sample estimates:
## mean in group lleuger    mean in group pesat
##          433.90          636.65
```

El valor p és molt inferior als nivells 0.05 amb un 95% de nivell de confiança, per tant, podem rebutjar la hipòtesis nul·la del qual el total de punts del Pokemons, són igual amb el pes lleuger i pesat. Es pot dir que els Pokemons amb més pes tenen un total\_points superior.

### Test de dues mostres de Pokemons de tipus Poison Segons pes:

H0:  $\mu_{\text{lleuger}} = \mu_{\text{pesat}}$  H1:  $\mu_{\text{lleuger}} < \mu_{\text{pesat}}$

Utilitzarem un test de dues mostres independents. Pel teorema del límit central, assumim normalitat.

```
var.test(pokedexAnalysis.Poison$total_points[pokedexAnalysis.Poison$Fweight=="lleuger"], pokedexAnalysis.Poison$total_points[pokedexAnalysis.Poison$Fweight=="pesat"], pokedexAnalysis.Poison$total_points, data.names=c("lleuger", "pesat"))

##
## F test to compare two variances
##
## data:  pokedexAnalysis.Poison$total_points[pokedexAnalysis.Poison$Fweight == "lleuger"] and pokedexAnalysis.Poison$total_points[pokedexAnalysis.Poison$Fweight == "pesat"]
## F = 0.1432, num df = 31, denom df = 6, p-value = 9.145e-05
## alternative hypothesis: true ratio of variances is less than 1
## 95 percent confidence interval:
##  0.0000000 0.3450336
## sample estimates:
## ratio of variances
##          0.1432012
```

El resultat del test F és que podem assumir igualtat de variàncies. Per tant, apliquem test t de dues mostres independents per a la diferència de mitjanes, variàncies desconegudes i iguals. El test és unilateral.

```
t.test(total_points~Fweight_kg,pokedexAnalysis.Poison,var.equal=T, alternative="less")

##
## Two Sample t-test
##
## data:  total_points by Fweight_kg
## t = -4.3142, df = 37, p-value = 5.724e-05
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##      -Inf -137.4004
## sample estimates:
## mean in group lleuger    mean in group pesat
##          392.2188          617.8571
```

El valor p és molt inferior als nivells 0.05 amb un 95% de nivell de confiança, per tant, podem rebutjar la hipòtesis nul·la del qual el total de punts dels Pokemons, són igual amb el pes lleuger i pesat. Es pot dir que els Pokemons amb més pes tenen un total\_points superior.

### Test de dues mostres de Pokemons de tipus Steel Segons pes:

$H_0: \mu_{lleuger} = \mu_{pesat}$   $H_1: \mu_{lleuger} < \mu_{pesat}$

Utilitzarem un test de dues mostres independents. Pel teorema del límit central, assumim normalitat.

```
var.test(pokedexAnalysis.Steel$total_points[pokedexAnalysis.Steel$Fweight=="lleuger"], pokedexAnalysis.Steel$total_points[pokedexAnalysis.Steel$Fweight=="pesat"], pokedexAnalysis.Steel$total_points, data.names=c("lleuger", "pesat"))

##
## F test to compare two variances
##
## data:  pokedexAnalysis.Steel$total_points[pokedexAnalysis.Steel$Fweight == "lleuger"] and pokedexAnalysis.Steel$total_points[pokedexAnalysis.Steel$Fweight == "pesat"]
## F = 0.84335, num df = 16, denom df = 18, p-value = 0.3688
## alternative hypothesis: true ratio of variances is less than 1
## 95 percent confidence interval:
##  0.0000000 1.941085
## sample estimates:
## ratio of variances
##          0.84335
```

El resultat del test F és que podem assumir igualtat de variàncies. Per tant, apliquem test t de dues mostres independents per a la diferència de mitjanes, variàncies desconegudes i iguals. El test és unilateral.

```
t.test(total_points~Fweight_kg,pokedexAnalysis.Steel,var.equal=T, alternative="less")
```

```
##
## Two Sample t-test
##
## data: total_points by Fweight_kg
## t = -3.3817, df = 34, p-value = 0.0009125
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##      -Inf -58.06748
## sample estimates:
## mean in group lleuger    mean in group pesat
##           420.1765           536.3158
```

El valor p és molt inferior als nivells 0.05 amb un 95% de nivell de confiança, per tant, podem rebutjar la hipòtesis nul·la del qual el total de punts del Pokemons, són igual amb el pes lleuger i pesat. Es pot dir que els Pokemons amb més pes tenen un total\_points superior.

#### Test de dues mostres de Pokemons de tipus Psychic Segons pes:

$H_0: \mu_{lleuger} = \mu_{pesat}$   $H_1: \mu_{lleuger} < \mu_{pesat}$

Utilitzarem un test de dues mostres independents. Pel teorema del límit central, assumim normalitat.

```
var.test(pokedexAnalysis.Psychic$total_points[pokedexAnalysis.Psychic$Fweight=="lleuger"], pokedexAnaly
```

```
##
## F test to compare two variances
##
## data: pokedexAnalysis.Psychic$total_points[pokedexAnalysis.Psychic$Fweight == "lleuger"] and pokedexAnalysis.Psychic$total_points[pokedexAnalysis.Psychic$Fweight == "pesat"]
## F = 1.289, num df = 60, denom df = 14, p-value = 0.6902
## alternative hypothesis: true ratio of variances is less than 1
## 95 percent confidence interval:
##  0.000000 2.397913
## sample estimates:
## ratio of variances
##           1.289033
```

El resultat del test F és que podem assumir igualtat de variàncies. Per tant, apliquem test t de dues mostres independents per a la diferència de mitjanes, variàncies desconegudes i iguals. El test és unilateral.

```
t.test(total_points~Fweight_kg,pokedexAnalysis.Psychic,var.equal=T, alternative="less")
```

```
##
## Two Sample t-test
##
## data: total_points by Fweight_kg
## t = -4.8174, df = 74, p-value = 3.766e-06
## alternative hypothesis: true difference in means is less than 0
```



```
## 95 percent confidence interval:
##      -Inf -112.3364
## sample estimates:
## mean in group lleuger    mean in group pesat
##           451.4262           623.1333
```

El valor p és molt inferior als nivells 0.05 amb un 95% de nivell de confiança, per tant, podem rebutjar la hipòtesis nul·la del qual el total de punts del Pokemons, són igual amb el pes lleuger i pesat. Es pot dir que els Pokemons amb més pes tenen un total\_points superior.

## Test de dues mostres: el total de punts (força) del Pokémon en funció l'alçada

### Test de dues mostres de Tots els Pokemons Segons alçada:

$H_0: \mu_{\text{baix}} = \mu_{\text{alt}}$   $H_1: \mu_{\text{baix}} < \mu_{\text{alt}}$

Utilitzarem un test de dues mostres independents. Pel teorema del límit central, assumim normalitat.

```
var.test(pokedexAnalysis$total_points[pokedexAnalysis$Fheight_m=="baix"], pokedexAnalysis$total_points[

##
## F test to compare two variances
##
## data:  pokedexAnalysis$total_points[pokedexAnalysis$Fheight_m == "baix"] and pokedexAnalysis$total_p
## F = 1.2677, num df = 687, denom df = 339, p-value = 0.9934
## alternative hypothesis: true ratio of variances is less than 1
## 95 percent confidence interval:
##  0.00000 1.47619
## sample estimates:
## ratio of variances
##           1.267745
```

El resultat del test F és que podem assumir igualtat de variàncies. Per tant, apliquem test t de dues mostres independents per a la diferència de mitjanes, variàncies desconegudes i iguals. El test és unilateral.

```
t.test(total_points~Fheight_m,pokedexAnalysis,var.equal=T, alternative="less")

##
## Two Sample t-test
##
## data:  total_points by Fheight_m
## t = 25.802, df = 1026, p-value = 1
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##      -Inf 172.4893
## sample estimates:
## mean in group alt mean in group baix
##           546.0882           383.9448
```

El valor p és superior als nivells 0.05 amb un 95% de nivell de confiança, per tant, podem acceptar la hipòtesis nul·la del qual el total de punts del Pokemons, són igual amb l'alçada alt i baix.

### Test de dues mostres de Pokemons Llegendaris Segons alçada:

H0:  $\mu_{\text{baix}} = \mu_{\text{alt}}$  H1:  $\mu_{\text{baix}} < \mu_{\text{alt}}$

Utilitzarem un test de dues mostres independents. Pel teorema del límit central, assumim normalitat.

```
var.test(pokedexAnalysis.Legendary$total_points[pokedexAnalysis.Legendary$Fheight_m=="baix"], pokedexAna

##
## F test to compare two variances
##
## data:  pokedexAnalysis.Legendary$total_points[pokedexAnalysis.Legendary$Fheight_m ==  and pokedexAna
## F = 3.0912, num df = 2, denom df = 35, p-value = 0.942
## alternative hypothesis: true ratio of variances is less than 1
## 95 percent confidence interval:
##  0.0000 60.1765
## sample estimates:
## ratio of variances
##          3.091179
```

El resultat del test F és que podem assumir igualtat de variàncies. Per tant, apliquem test t de dues mostres independents per a la diferència de mitjanes, variàncies desconegudes i iguals. El test és unilateral.

```
t.test(total_points~Fheight_m,pokedexAnalysis.Legendary,var.equal=T, alternative="less")

##
## Two Sample t-test
##
## data:  total_points by Fheight_m
## t = 6.5019, df = 37, p-value = 1
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##      -Inf 433.2955
## sample estimates:
## mean in group alt mean in group baix
##          706.0278          362.0000
```

El valor p és superior als nivells 0.05 amb un 95% de nivell de confiança, per tant, podem acceptar la hipòtesis nul·la del qual el total de punts del Pokemons, són igual amb l'alçada alt i baix.

### Test de dues mostres de Pokemons Sub Llegendaris Segons alçada:

H0:  $\mu_{\text{baix}} = \mu_{\text{alt}}$  H1:  $\mu_{\text{baix}} < \mu_{\text{alt}}$

Utilitzarem un test de dues mostres independents. Pel teorema del límit central, assumim normalitat.

```
var.test(pokedexAnalysis.SubLegendary$total_points[pokedexAnalysis.SubLegendary$Fheight_m=="baix"], pok

##
## F test to compare two variances
##
## data:  pokedexAnalysis.SubLegendary$total_points[pokedexAnalysis.SubLegendary$Fheight_m ==  and poked
## F = 2.3771, num df = 8, denom df = 35, p-value = 0.9632
```

```
## alternative hypothesis: true ratio of variances is less than 1
## 95 percent confidence interval:
##  0.000000 7.270659
## sample estimates:
## ratio of variances
##          2.377138
```

El resultat del test F és que podem assumir igualtat de variàncies. Per tant, apliquem test t de dues mostres independents per a la diferència de mitjanes, variàncies desconegudes i iguals. El test és unilateral.

```
t.test(total_points~Fheight_m,pokedexAnalysis.SubLegendary,var.equal=T, alternative="less")
```

```
##
## Two Sample t-test
##
## data: total_points by Fheight_m
## t = 1.8515, df = 43, p-value = 0.9645
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##      -Inf 51.09122
## sample estimates:
## mean in group alt mean in group baix
##          586.7778          560.0000
```

El valor p és superior als nivells 0.05 amb un 95% de nivell de confiança, per tant, podem acceptar la hipòtesis nul·la del qual el total de punts del Pokemons, són igual amb l'alçada alt i baix.

### Test de dues mostres de Pokemons de tipus Normal Segons alçada:

$H_0: \mu_{\text{baix}} = \mu_{\text{alt}}$   $H_1: \mu_{\text{baix}} < \mu_{\text{alt}}$

Utilitzarem un test de dues mostres independents. Pel teorema del límit central, assumim normalitat.

```
var.test(pokedexAnalysis.Normal$total_points[pokedexAnalysis.Normal$Fheight_m=="baix"], pokedexAnalysis.
```

```
##
## F test to compare two variances
##
## data: pokedexAnalysis.Normal$total_points[pokedexAnalysis.Normal$Fheight_m == "baix"] and pokedexAnalysis.Normal$total_points[pokedexAnalysis.Normal$Fheight_m == "alt"]
## F = 2.1678, num df = 659, denom df = 254, p-value = 0.0321
## alternative hypothesis: true ratio of variances is less than 1
## 95 percent confidence interval:
##  0.00000 2.56503
## sample estimates:
## ratio of variances
##          2.167811
```

El resultat del test F és que podem assumir igualtat de variàncies. Per tant, apliquem test t de dues mostres independents per a la diferència de mitjanes, variàncies desconegudes i iguals. El test és unilateral.

```
t.test(total_points~Fheight_m,pokedexAnalysis.Normal,var.equal=T, alternative="less")
```

```
##
## Two Sample t-test
##
## data: total_points by Fheight_m
## t = 22.069, df = 913, p-value = 1
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##      -Inf 147.7555
## sample estimates:
## mean in group alt mean in group baix
##      514.3882      376.8909
```

El valor p és superior als nivells 0.05 amb un 95% de nivell de confiança, per tant, podem acceptar la hipòtesis nul·la del qual el total de punts del Pokemons, són igual amb l'alçada alt i baix.

### Test de dues mostres de Pokemons de tipus Bug Segons alçada:

$H_0: \mu_{\text{baix}} = \mu_{\text{alt}}$   $H_1: \mu_{\text{baix}} < \mu_{\text{alt}}$

Utilitzarem un test de dues mostres independents. Pel teorema del límit central, assumim normalitat.

```
var.test(pokedexAnalysis.Bug$total_points[pokedexAnalysis.Bug$Fheight_m=="baix"], pokedexAnalysis.Bug$total_points[pokedexAnalysis.Bug$Fheight_m=="alt"], var.equal=T, alternative="less")
```

```
##
## F test to compare two variances
##
## data: pokedexAnalysis.Bug$total_points[pokedexAnalysis.Bug$Fheight_m=="baix"] and pokedexAnalysis.Bug$total_points[pokedexAnalysis.Bug$Fheight_m=="alt"]
## F = 3.1787, num df = 61, denom df = 18, p-value = 0.9958
## alternative hypothesis: true ratio of variances is less than 1
## 95 percent confidence interval:
##  0.000000 5.643945
## sample estimates:
## ratio of variances
##      3.178731
```

El resultat del test F és que podem assumir igualtat de variàncies. Per tant, apliquem test t de dues mostres independents per a la diferència de mitjanes, variàncies desconegudes i iguals. El test és unilateral.

```
t.test(total_points~Fheight_m,pokedexAnalysis.Bug,var.equal=T, alternative="less")
```

```
##
## Two Sample t-test
##
## data: total_points by Fheight_m
## t = 7.4047, df = 79, p-value = 1
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##      -Inf 221.6795
## sample estimates:
## mean in group alt mean in group baix
##      523.1579      342.1613
```

El valor p és superior als nivells 0.05 amb un 95% de nivell de confiança, per tant, podem acceptar la hipòtesis nul·la del qual el total de punts del Pokemons, són igual amb l'alçada alt i baix.

### Test de dues mostres de Pokemons de tipus Dragon Segons alçada:

H0:  $\mu_{\text{baix}} = \mu_{\text{alt}}$  H1:  $\mu_{\text{baix}} < \mu_{\text{alt}}$

Utilitzarem un test de dues mostres independents. Pel teorema del límit central, assumim normalitat.

```
var.test(pokedexAnalysis.Dragon$total_points[pokedexAnalysis.Dragon$Fheight_m=="baix"], pokedexAnalysis.Dragon$total_points[pokedexAnalysis.Dragon$Fheight_m=="alt"], pokedexAnalysis.Dragon$total_points[pokedexAnalysis.Dragon$Fheight_m=="alt"], data.names=c("baix", "alt"), var.equal=T, alternative="less")

##
## F test to compare two variances
##
## data:  pokedexAnalysis.Dragon$total_points[pokedexAnalysis.Dragon$Fheight_m == "baix"] and pokedexAnalysis.Dragon$total_points[pokedexAnalysis.Dragon$Fheight_m == "alt"]
## F = 0.53222, num df = 11, denom df = 27, p-value = 0.1359
## alternative hypothesis: true ratio of variances is less than 1
## 95 percent confidence interval:
##  0.000000 1.377238
## sample estimates:
## ratio of variances
##          0.5322201
```

El resultat del test F és que podem assumir igualtat de variàncies. Per tant, apliquem test t de dues mostres independents per a la diferència de mitjanes, variàncies desconegudes i iguals. El test és unilateral.

```
t.test(total_points~Fheight_m,pokedexAnalysis.Dragon,var.equal=T, alternative="less")

##
## Two Sample t-test
##
## data:  total_points by Fheight_m
## t = 6.488, df = 38, p-value = 1e-06
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##      -Inf 293.2617
## sample estimates:
## mean in group alt mean in group baix
##          605.1071          372.3333
```

El valor p és superior als nivells 0.05 amb un 95% de nivell de confiança, per tant, podem acceptar la hipòtesis nul·la del qual el total de punts del Pokemons, són igual amb l'alçada alt i baix.

### Test de dues mostres de Pokemons de tipus Poison Segons alçada:

H0:  $\mu_{\text{baix}} = \mu_{\text{alt}}$  H1:  $\mu_{\text{baix}} < \mu_{\text{alt}}$

Utilitzarem un test de dues mostres independents. Pel teorema del límit central, assumim normalitat.

```
var.test(pokedexAnalysis.Poison$total_points[pokedexAnalysis.Poison$Fheight_m=="baix"], pokedexAnalysis.Poison$total_points[pokedexAnalysis.Poison$Fheight_m=="alt"], pokedexAnalysis.Poison$total_points[pokedexAnalysis.Poison$Fheight_m=="alt"], data.names=c("baix", "alt"), var.equal=T, alternative="less")

##
## F test to compare two variances
##
## data:  pokedexAnalysis.Poison$total_points[pokedexAnalysis.Poison$Fheight_m == "baix"] and pokedexAnalysis.Poison$total_points[pokedexAnalysis.Poison$Fheight_m == "alt"]
## F = 0.20543, num df = 25, denom df = 12, p-value = 0.0004168
```

```
## alternative hypothesis: true ratio of variances is less than 1
## 95 percent confidence interval:
##  0.0000000 0.4447285
## sample estimates:
## ratio of variances
##           0.2054276
```

El resultat del test F és que podem assumir igualtat de variàncies. Per tant, apliquem test t de dues mostres independents per a la diferència de mitjanes, variàncies desconegudes i iguals. El test és unilateral.

```
t.test(total_points~Fweight_kg,pokedexAnalysis.Poison,var.equal=T, alternative="less")
```

```
##
## Two Sample t-test
##
## data: total_points by Fweight_kg
## t = -4.3142, df = 37, p-value = 5.724e-05
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##      -Inf -137.4004
## sample estimates:
## mean in group lleuger mean in group pesat
##           392.2188           617.8571
```

El valor p és molt inferior als nivells 0.05 amb un 95% de nivell de confiança, per tant, podem rebutjar la hipòtesis nul·la del qual el total de punts del Pokemons, són igual els alts i baixos. Es pot dir que els Pokemons amb més alçada tenen un total\_points superior.

### Test de dues mostres de Pokemons de tipus Steel Segons alçada:

$H_0: \mu_{\text{baix}} = \mu_{\text{alt}}$   $H_1: \mu_{\text{baix}} < \mu_{\text{alt}}$

Utilitzarem un test de dues mostres independents. Pel teorema del límit central, assumim normalitat.

```
var.test(pokedexAnalysis.Steel$total_points[pokedexAnalysis.Steel$Fheight_m=="baix"], pokedexAnalysis.St
```

```
##
## F test to compare two variances
##
## data: pokedexAnalysis.Steel$total_points[pokedexAnalysis.Steel$Fheight_m == "baix"] and pokedexAnalysis.St
## F = 2.3403, num df = 19, denom df = 15, p-value = 0.95
## alternative hypothesis: true ratio of variances is less than 1
## 95 percent confidence interval:
##  0.000000 5.228398
## sample estimates:
## ratio of variances
##           2.340309
```

El resultat del test F és que podem assumir igualtat de variàncies. Per tant, apliquem test t de dues mostres independents per a la diferència de mitjanes, variàncies desconegudes i iguals. El test és unilateral.

```
t.test(total_points~Fheight_m,pokedexAnalysis.Steel,var.equal=T, alternative="less")
```

```
##
## Two Sample t-test
##
## data: total_points by Fheight_m
## t = 5.5584, df = 34, p-value = 1
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##      -Inf 209.2936
## sample estimates:
## mean in group alt mean in group baix
##      570.625      410.150
```

El valor p és superior als nivells 0.05 amb un 95% de nivell de confiança, per tant, podem acceptar la hipòtesis nul·la del qual el total de punts del Pokemons, són igual amb l'alçada alt i baix.

### Test de dues mostres de Pokemons de tipus Psychic Segons alçada:

$H_0: \mu_{baix} = \mu_{alt}$   $H_1: \mu_{baix} < \mu_{alt}$

Utilitzarem un test de dues mostres independents. Pel teorema del límit central, assumim normalitat.

```
var.test(pokedexAnalysis.Psychic$total_points[pokedexAnalysis.Psychic$Fheight_m=="baix"], pokedexAnalysis
```

```
##
## F test to compare two variances
##
## data: pokedexAnalysis.Psychic$total_points[pokedexAnalysis.Psychic$Fheight_m == "baix"] and pokedexAnalysis.Psychic$total_points[pokedexAnalysis.Psychic$Fheight_m == "alt"]
## F = 1.485, num df = 46, denom df = 28, p-value = 0.8664
## alternative hypothesis: true ratio of variances is less than 1
## 95 percent confidence interval:
##  0.000000 2.556441
## sample estimates:
## ratio of variances
##      1.485033
```

El resultat del test F és que podem assumir igualtat de variàncies. Per tant, apliquem test t de dues mostres independents per a la diferència de mitjanes, variàncies desconegudes i iguals. El test és unilateral.

```
t.test(total_points~Fheight_m,pokedexAnalysis.Psychic,var.equal=T, alternative="less")
```

```
##
## Two Sample t-test
##
## data: total_points by Fheight_m
## t = 7.1286, df = 74, p-value = 1
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##      -Inf 226.6471
## sample estimates:
## mean in group alt mean in group baix
##      598.9310      415.2128
```

El valor p és superior als nivells 0.05 amb un 95% de nivell de confiança, per tant, podem acceptar la hipotesis nul·la del qual el total de punts del Pokemons, són igual amb l'alçada alt i baix.

**4.3.3 Fer una predicció de com les variables d'estadística base, el pes i l'alçada influeixen amb la variable base\_experience.** Aplicarem cinc models de regressió lineal múltiple que cadascun d'ells usaran diferents variables explicatives: attack, defense, sp\_attack, sp\_defense, hp, speed, Fheight\_m i Fweight\_kg.

Especificarem en el nivell base (en el relevel): per la variable Fheight\_m, la categoria "alt", i per la variable Fweight\_kg la categoria "lleuger".

```
heightR <- relevel(pokedexAnalysis$Fheight_m, ref="alt" )

weightR <- relevel(pokedexAnalysis$Fweight_kg, ref="lleuger")

pokedexAnalysis$heightR <- heightR
pokedexAnalysis$weightR <- weightR

mod1 <- lm(base_experience~attack+defense+sp_attack+sp_defense+hp+speed+heightR+weightR, pokedexAnalysis)
mod2 <- lm(base_experience~attack+defense+sp_attack+sp_defense+hp+speed, pokedexAnalysis )
mod3 <- lm(base_experience~attack+defense+hp+speed+heightR+weightR, pokedexAnalysis )
mod4 <- lm(base_experience~attack+sp_attack+hp+speed+heightR+weightR, pokedexAnalysis )
mod5 <- lm(base_experience~attack+defense+sp_attack+sp_defense+heightR+weightR, pokedexAnalysis )
```

Per els anteriors models de regressió lineal múltiple obtinguts, podem utilitzar el coeficient de determinació per mesurar la bondat d'ajust i quedar-nos amb el model que tingui millor coeficient.

```
#summary(mod1)
tabla.coeficientes <- matrix(c(1, summary(mod1)$r.squared, 2, summary(mod2)$r.squared, 3, summary(mod3)$r.squared, 4, summary(mod4)$r.squared, 5, summary(mod5)$r.squared),
                             nrow=5, byrow=TRUE,
                             dimnames=list(c("Model", "R^2"), 1:5))
tabla.coeficientes
```

```
##      Model      R^2
## [1,]      1 0.8806787
## [2,]      2 0.8776792
## [3,]      3 0.7898483
## [4,]      4 0.7837578
## [5,]      5 0.7882521
```

El primer model és el més convenient ja que té un major coeficient de determinació. Ara utilitzarem aquest model per fer prediccions amb diferents valors:

```
F0<-data.frame( attack=130, defense=80, sp_attack=110, sp_defense=90, hp=75,speed=20,heightR="alt",weightR="lleuger")
predict(mod1, F0)
```

```
##      1
## 185.7625
```



```
F1<-data.frame( attack=60, defense=57, sp_attack=101, sp_defense=86, hp=100,speed=180,heightR="baix",weightR="normal")
predict(mod1, F1)
```

```
##          1
## 246.5225
```

```
F2<-data.frame( attack=49, defense=49, sp_attack=65, sp_defense=65, hp=45,speed=45,heightR="baix",weightR="normal")
predict(mod1, F2)
```

```
##          1
## 79.39432
```

```
F3<-data.frame( attack=130, defense=80, sp_attack=110, sp_defense=90, hp=75,speed=20,heightR="alt",weightR="normal")
predict(mod1, F3)
```

```
##          1
## 196.4184
```

```
##valors minims
```

```
F4<-data.frame( attack=5, defense=5, sp_attack=10, sp_defense=20, hp=1,speed=5,heightR="baix",weightR="normal")
predict(mod1, F4)
```

```
##          1
## -89.07218
```

```
##valors mitjans
```

```
F5<-data.frame( attack=80.12, defense=74.48, sp_attack=71.73, sp_defense=72.13, hp=69.58,speed=68.53,heightR="baix",weightR="normal")
predict(mod1, F5)
```

```
##          1
## 151.8014
```

```
##valors màxims
```

```
F6<-data.frame( attack=190, defense=250, sp_attack=194, sp_defense=250, hp=255,speed=180,heightR="baix",weightR="normal")
predict(mod1, F6)
```

```
##          1
## 721.0123
```

En les anteriors prediccions del model, es pot observar el base\_experience. La variable Hp, es una de les variables que té més pes alhora que tingui més base\_experience. S'ha fet un anàlisi per veure el comportament segons els valors mínims, mitjans i màxims, i s'ha observat que amb valors mínims fa una predicció d'un Pokémon inexistent, ja que no hi ha cap Pokémon amb una base\_experience negativa. Quan és valors mitjans, s'adapta bastant amb la mitjana de base\_experience que és de 153.8. I quan es tracta amb valors màxims, deixa una base\_experience, més elevada que la que podem trobar en ella.

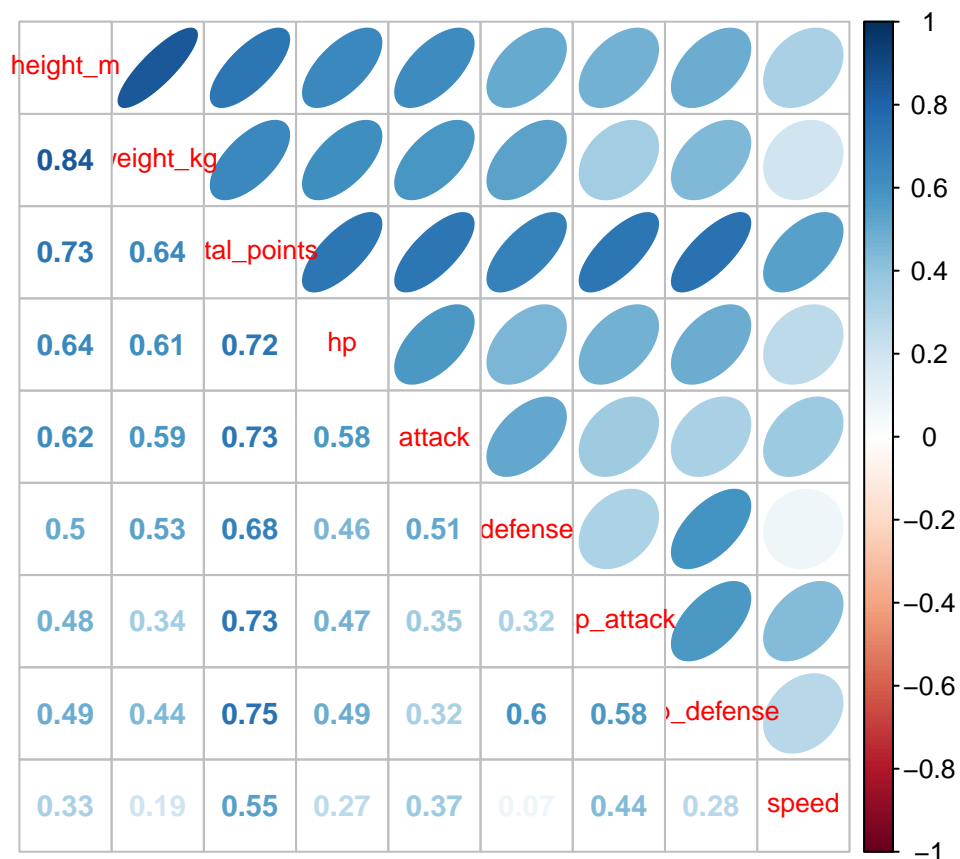
## 5. Representació dels resultats a partir de taules i gràfiques.

Com s'ha pogut veure, durant l'anàlisi s'han realitzat taules i gràfiques per fer un anàlisi visual de les dades.

A continuació mostrem les correlacions realitzades mitjançant gràfics:

```
corrplot.mixed(corr.general,upper="ellipse",number.cex=.9,tl.cex=.8)
```

### Correlació dels Pokemons en general

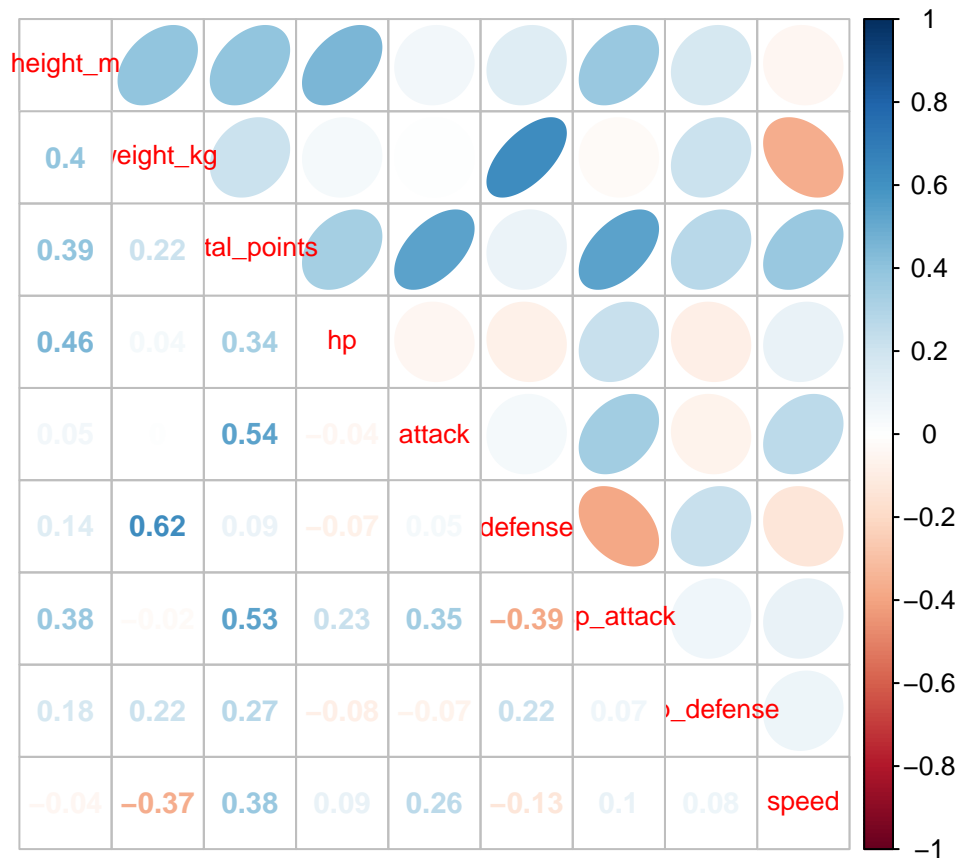


En el gràfic, es pot observar d'una manera ràpida, que no hi ha correlacions negatives. Quan és un el·lipse més pronunciat, té un color blau més fosc que indica més correlació. Quan més arrodonit i color més claret, menys hi ha correlació positiva.

Es pot observar que la correlació més forta quan s'estudien els Pokemons en general, és el seu pes i la seva alçada.

```
corrplot.mixed(corr.legendary,upper="ellipse",number.cex=.9,tl.cex=.8)
```

### Correlació dels Pokemons Legendaris

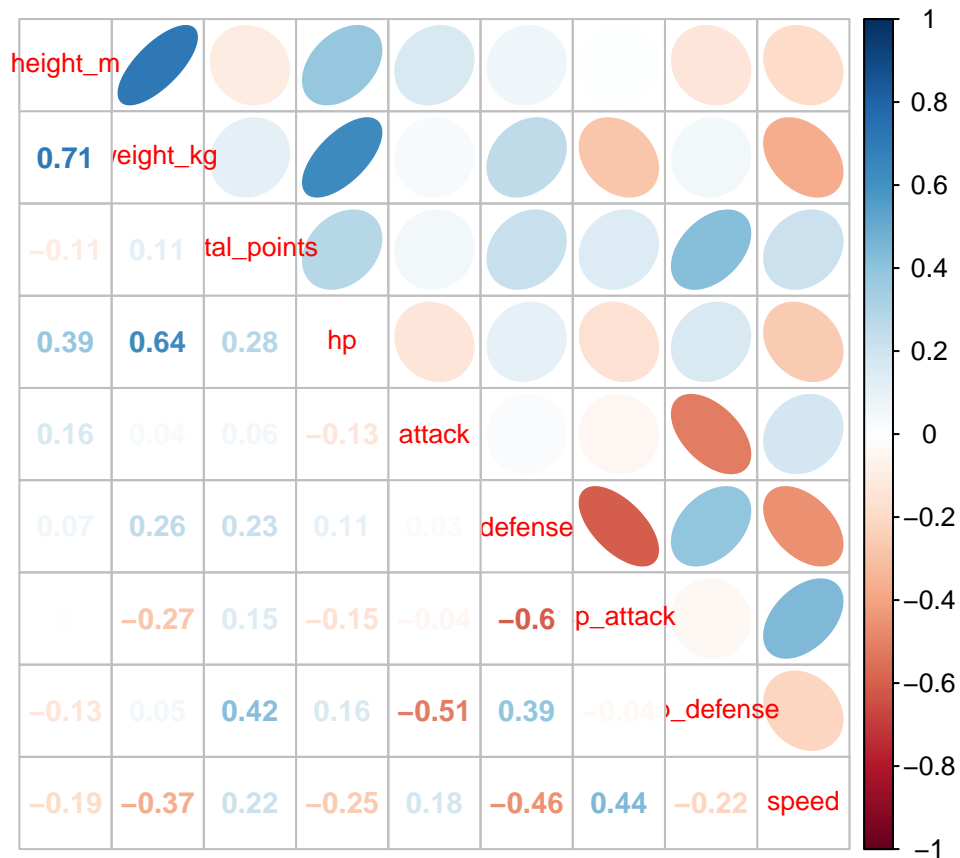


En el gràfic, es pot observar d'una manera ràpida, que hi ha poques correlacions negatives. Quan és un el·lipse més pronunciat, té un color blau més fosc que indica més correlació. Quan més arrodonit i color més claret, menys hi ha correlació positiva. Les figures transparents ens indiquen que la correlació està pels voltants del 0. I les figures de color ataronjat ens indiquen que hi ha un mica de correlació negativa.

Es pot observar que la correlació més forta quan s'estudien els Pokemons Llegendaris, és el seu pes i la seva defensa. I més negativa és defense i sp\_attack, que vol dir que quan una d'elles puja l'altre disminueix.

```
corrplot.mixed(corr.sublegendary,upper="ellipse",number.cex=.9,tl.cex=.8)
```

## Correlació dels Pokemons Sub Legendaris

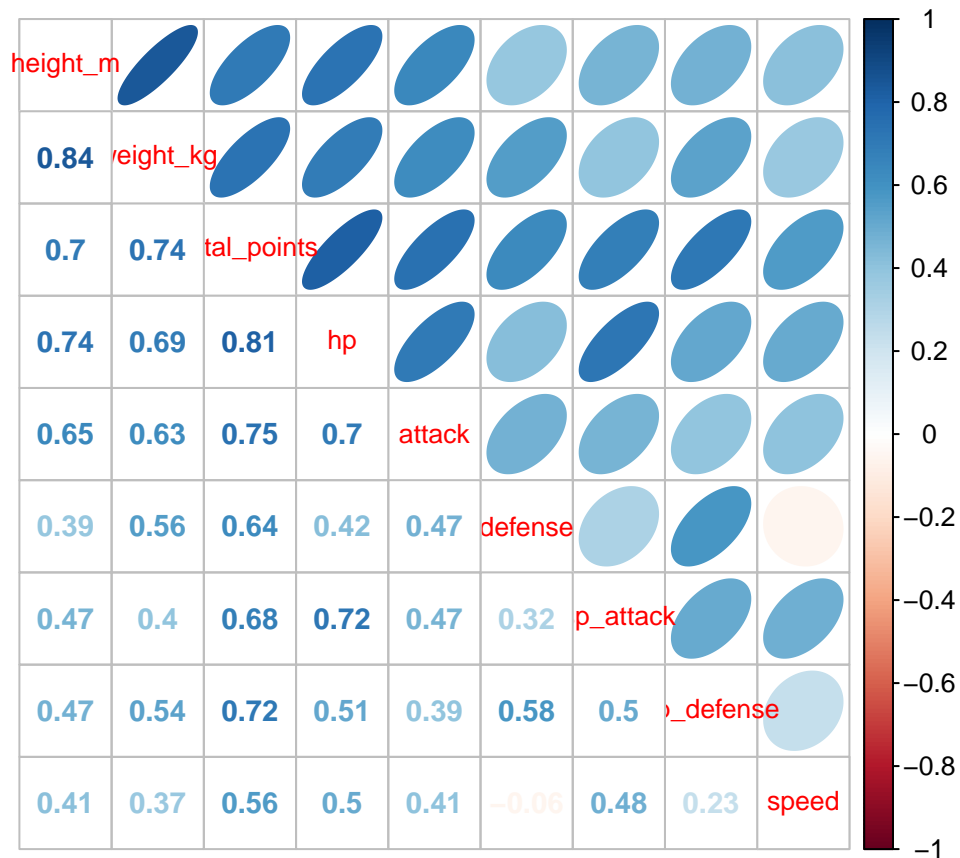


En el gràfic, es pot observar d'una manera ràpida, que hi ha bastantes correlacions negatives. Quan és un el·lipse més pronunciat, té un color blau més fosc que indica més correlació. Quan més arrodonit i color més claret, menys hi ha correlació positiva. Les figures transparents ens indiquen que la correlació està pels voltants del 0. I les figures de color ataronjat ens indiquen que hi ha un mica de correlació negativa, i les figures de color més vermellós, ens indiquen que hi ha molta correlació negativa.

Es pot observar que la correlació més forta quan s'estudien els Pokemons Sub Llegendaris, és el seu pes i la seva l'alçada, i més negativa és defense i sp\_attack, que vol dir que quan una d'elles puja l'altre disminueix.

```
corrplot.mixed(corr.bug,upper="ellipse",number.cex=.9,tl.cex=.8)
```

### Correlació dels Pokemons de tipus Bug

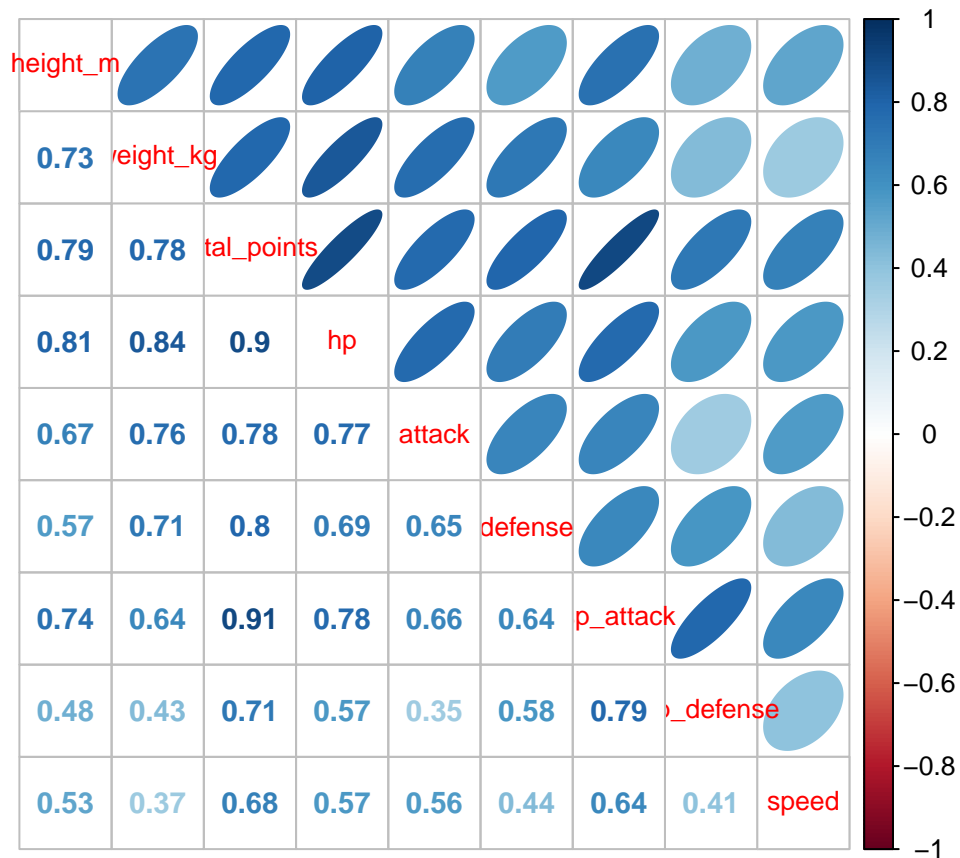


En el gràfic, es pot observar d'una manera ràpida, que no hi ha correlacions negatives. Quan és un el·lipse més pronunciat, té un color blau més fosc que indica més correlació. Quan més arrodonit i color més claret, menys hi ha correlació positiva.

Es pot observar que la correlació més forta quan s'estudien els Pokemons de tipus bug, és el seu pes i la seva alçada.

```
corrplot.mixed(corr.dragon,upper="ellipse",number.cex=.9,tl.cex=.8)
```

**Correlació dels Pokemons de tipus Dragon**

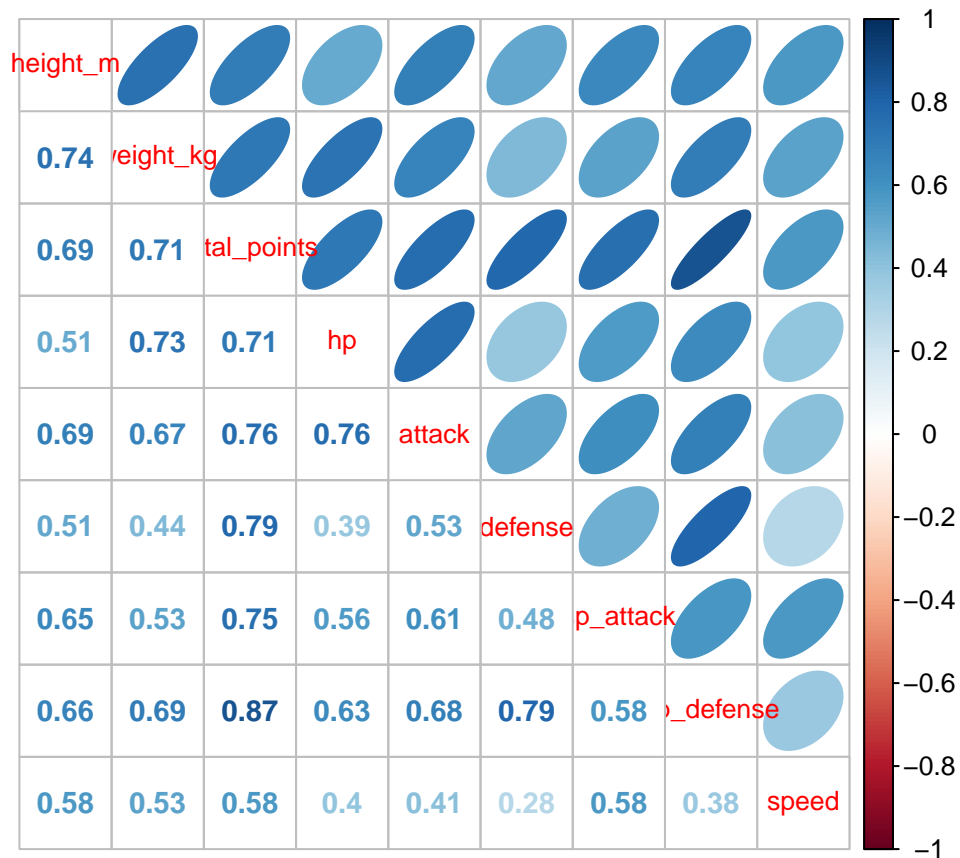


En el gràfic, es pot observar d'una manera ràpida, que no hi ha correlacions negatives. Quan és un el·lipse més pronunciat, té un color blau més fosc que indica més correlació. Quan més arrodonit i color més claret, menys hi ha correlació positiva.

Es pot observar que la correlació més forta quan s'estudien els Pokemons de tipus dragon, és el seu total\_points i la sp\_attack.

```
corrplot.mixed(corr.poison,upper="ellipse",number.cex=.9,tl.cex=.8)
```

## Correlació dels Pokemons de tipus Poison

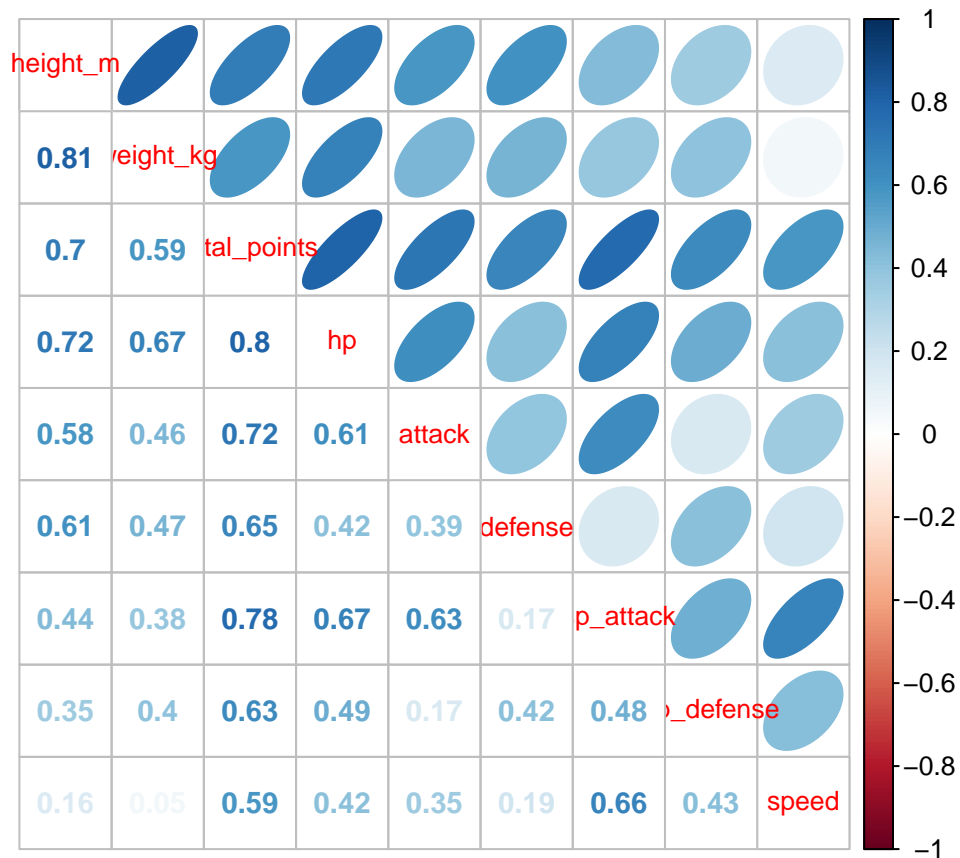


En el gràfic, es pot observar d'una manera ràpida, que no hi ha correlacions negatives. Quan és un el·lipse més pronunciat, té un color blau més fosc que indica més correlació. Quan més arrodonit i color més claret, menys hi ha correlació positiva.

Es pot observar que la correlació més forta quan s'estudien els Pokemons de tipus Poison, és el seu total\_points i la sp\_defense.

```
corrplot.mixed(corr.steel,upper="ellipse",number.cex=.9,tl.cex=.8)
```

### Correlació dels Pokemons de tipus Steel



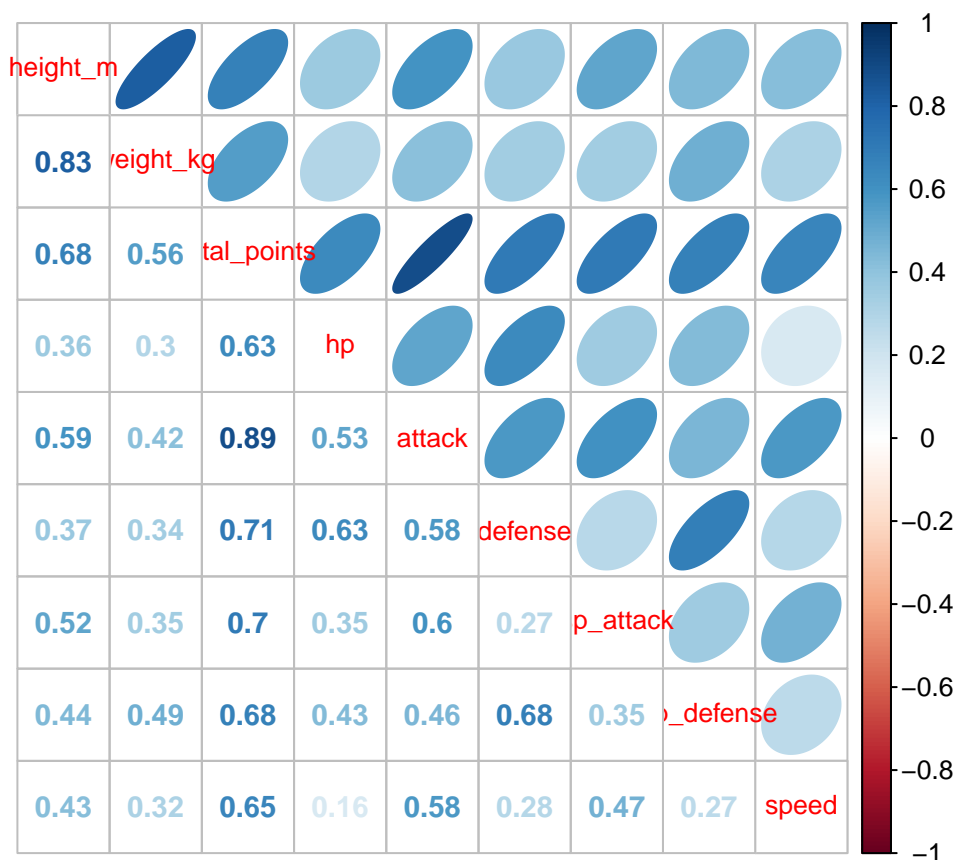
En el gràfic, es pot observar d'una manera ràpida, que no hi ha correlacions negatives. Quan és un el·lipse més pronunciat, té un color blau més fosc que indica més correlació. Quan més arrodonit i color més claret, menys hi ha correlació positiva.

Es pot observar que la correlació més forta quan s'estudien els Pokemons de tipus Steel, és el seu pes i la l'alçada.

```
corrplot.mixed(corr.psychic,upper="ellipse",number.cex=.9,tl.cex=.8)
```

Correlació dels Pokemons de tipus Psychic





En el gràfic, es pot observar d'una manera ràpida, que no hi ha correlacions negatives. Quan és un el·lipse més pronunciat, té un color blau més fosc que indica més correlació. Quan més arrodonit i color més claret, menys hi ha correlació positiva.

Es pot observar que la correlació més forta quan s'estudien els Pokemons de tipus Steel, és el total\_points i l'attack.

## 6. Resolució del problema. A partir dels resultats obtinguts, quines són les conclusions? Els resultats permeten respondre al problema?

Amb l'estudi realitzat, hem estudiat especialment variables del Pokémon com la seva alçada i el seu pes, juntament amb les variables d'estadístiques bàsiques: total de punts, HP, atac, defensa, velocitat, atac especial i defensa especial i llavors de la informació d'entrenament: l'experiència base. Pensem que aquestes variables analitzades ens han destapat coneixaments que no ens esperàvem, com que els Pokémons menys estables són els Llegendaris i Sub llegendaris, hem pogut veure que les variables entre elles tenen poca relació. Creiem que això pot ser degut a que aquets Pokemons tenen unes característiques molt variades. Hem estudiat els tipus de Pokémon: bug, dragon, poison, psychic i steel, perquè amb l'estudi visual de la distribució del tipus amb el total de punts, són els que ens han cridat més l'atenció, els dragon, els més forts, els bug els més dèbils, els poison, els més destacats en els llegendaris, els psychic i amb els steels, els que tenen els quartils i els bigotis del boxplot, més proporcionats. Hem pogut concloure, en forma general, el pes i l'alçada dels pokemons es correlacionen molt bé, cosa que creiem que és força lògica ja que normalment quan es té més alçada també es té més pes. S'ha demostrat que els Pokemons de tipus dragon són els més forts, i tenen les variables d'estadístiques base més correlacionades. Els Sub llegendaris, els que tenen les variables d'estadístiques base amb correlacions més negatives. Els de tipus poison i steel són bastant iguals en correlació de les variables d'estadístiques base. I els de tipus bug, els més fluixets.

En relació a l'alçada i el pes, si influeix amb les estadístiques base, en general quan més pes, més total de punts, excepte els de tipus sub llegendaris que no hi ha entre si pesen poc o molt. En canvi en relació a la seva alçada, en la majoria dels casos l'alçada no influeix si són baixos o alts en relació al total\_points, excepte els de tipus Poison que quan més alts més total\_points.

Referent a l'estudi de l'experiència base, es pot dir que el que influeix una mica més és l'hp, però en general totes les variables d'estadística base tenen més o menys un pes per donar el valor de l'experiència base. Ens ha sobtat que quan hem posat els mínims que hi ha a la base de dades, el valor ha estat negatiu.

Per tant, podem dir que hem pogut resoldre les inquietuds que teníem respecte els Pokemons més forts, els llegendaris i sub llegendaris. També hem pogut observar que cada tipologia de Pokemons, té alguna particularitat en especial respecte al total\_points, cada ún d'ells es correlaciona amb més força amb una variable diferent. El de tipus Psychic és l'attack amb un 0.888, el de tipus Steel és l'HP amb un 0.804, el de tipus Poison és l'sp\_defense amb un 0.866, el de tipus Dragon és sp\_attack amb un 0.907, el de tipus Bug és HP amb un 0.813, els sub llegendaris és sp\_defense amb un 0.42, els llegendaris és attack amb un 0.539 i els de tipus normal és sp\_defense amb un 0.729.

## 7. Codi: Cal adjuntar el codi, preferiblement en R, amb el que s'ha realitzat la neteja, anàlisi i representació de les dades. Si ho preferiu, també podeu treballar en Python.

El codi R, està inclòs en aquest mateix informe.

## 8. Bibliografia

Donem gràcies a Mario Tormo Romero per Dataset de <https://www.kaggle.com/mariotormo/complete-pokemon-dataset-updated-090420>

- [https://pokemon.fandom.com/es/wiki/Crianza\\_Pok%C3%A9mon](https://pokemon.fandom.com/es/wiki/Crianza_Pok%C3%A9mon)
- [https://es.wikipedia.org/wiki/Coeficiente\\_de\\_correlaci%C3%B3n\\_de\\_Spearman](https://es.wikipedia.org/wiki/Coeficiente_de_correlaci%C3%B3n_de_Spearman)
- [https://pokemon.fandom.com/es/wiki/Experiencia\\_base](https://pokemon.fandom.com/es/wiki/Experiencia_base)

## 9. Contribucions

- **Investigació prèvia:** David Navarro, Antoni Llussà
- **Redacció respostes:** David Navarro, Antoni Llussà
- **Desenvolupament codi:** David Navarro, Antoni Llussà