Homework 6

PSTAT 131/231

Contents

Tree-Based Models

Note: Fitting ensemble tree-based models can take a little while to run. Consider running your models outside of the .Rmd, storing the results, and loading them in your .Rmd to minimize time to knit.

Exercise 1

Read in the data and set things up as in Homework 5:

- Use `clean_names()`
- Filter out the rarer Pokémon types
- Convert `type_1` and `legendary` to factors

Do an initial split of the data; you can choose the percentage for splitting. Stratify on the outcome variable.

Fold the training set using v-fold cross-validation, with `v = 5`. Stratify on the outcome variable.

Set up a recipe to predict `type_1` with `legendary`, `generation`, `sp_atk`, `attack`, `speed`, `defense`, `hp`, and `sp_def`:

- Dummy-code `legendary` and `generation`;
- Center and scale all predictors.

```
pokemon <- read.csv(file = "~/Pokemon.csv")
head(pokemon)
```

```
##   X.                  Name Type.1 Type.2 Total HP Attack Defense Sp..Atk
## 1  1             Bulbasaur  Grass Poison   318 45     49      49      65
## 2  2               Ivysaur  Grass Poison   405 60     62      63      80
## 3  3              Venusaur  Grass Poison   525 80     82      83     100
## 4  3 VenusaurMega Venusaur  Grass Poison   625 80    100     123     122
## 5  4            Charmander   Fire          309 39     52      43      60
## 6  5            Charmeleon   Fire          405 58     64      58      80
##   Sp..Def Speed Generation Legendary
## 1      65    45          1     False
## 2      80    60          1     False
```

```
## 3     100    80          1     False
## 4     120    80          1     False
## 5      50    65          1     False
## 6      65    80          1     False
```

```
library(janitor)
pokemon <- pokemon %>% clean_names()

pokemon_filter <- pokemon[pokemon$type_1 %in% c("Bug", "Fire", "Grass", "Normal", "Water", "Psychic"),]

pokemon_filter
```

```
##       x                    name  type_1   type_2 total  hp attack defense
## 1     1               Bulbasaur   Grass   Poison   318  45     49      49
## 2     2                 Ivysaur   Grass   Poison   405  60     62      63
## 3     3                Venusaur   Grass   Poison   525  80     82      83
## 4     3    VenusaurMega Venusaur   Grass   Poison   625  80    100     123
## 5     4              Charmander    Fire            309  39     52      43
## 6     5              Charmeleon    Fire            405  58     64      58
## 7     6               Charizard    Fire   Flying   534  78     84      78
## 8     6 CharizardMega Charizard X    Fire   Dragon   634  78    130     111
## 9     6 CharizardMega Charizard Y    Fire   Flying   634  78    104      78
## 10    7                Squirtle   Water            314  44     48      65
## 11    8               Wartortle   Water            405  59     63      80
## 12    9               Blastoise   Water            530  79     83     100
## 13    9  BlastoiseMega Blastoise   Water            630  79    103     120
## 14   10                Caterpie     Bug            195  45     30      35
## 15   11                 Metapod     Bug            205  50     20      55
## 16   12              Butterfree     Bug   Flying   395  60     45      50
## 17   13                  Weedle     Bug   Poison   195  40     35      30
## 18   14                  Kakuna     Bug   Poison   205  45     25      50
## 19   15                 Beedrill     Bug   Poison   395  65     90      40
## 20   15    BeedrillMega Beedrill     Bug   Poison   495  65    150      40
## 21   16                   Pidgey  Normal   Flying   251  40     45      40
## 22   17                Pidgeotto  Normal   Flying   349  63     60      55
## 23   18                  Pidgeot  Normal   Flying   479  83     80      75
## 24   18      PidgeotMega Pidgeot  Normal   Flying   579  83     80      80
## 25   19                  Rattata  Normal            253  30     56      35
## 26   20                 Raticate  Normal            413  55     81      60
## 27   21                  Spearow  Normal   Flying   262  40     60      30
## 28   22                   Fearow  Normal   Flying   442  65     90      65
## 43   37                   Vulpix    Fire            299  38     41      40
## 44   38                Ninetales    Fire            505  73     76      75
## 45   39               Jigglypuff  Normal    Fairy   270 115     45      20
## 46   40               Wigglytuff  Normal    Fairy   435 140     70      45
## 49   43                   Oddish   Grass   Poison   320  45     50      55
## 50   44                    Gloom   Grass   Poison   395  60     65      70
## 51   45                Vileplume   Grass   Poison   490  75     80      85
## 52   46                    Paras     Bug    Grass   285  35     70      55
## 53   47                 Parasect     Bug    Grass   405  60     95      80
## 54   48                  Venonat     Bug   Poison   305  60     55      50
## 55   49                 Venomoth     Bug   Poison   450  70     65      60
## 58   52                   Meowth  Normal            290  40     45      35
## 59   53                   Persian  Normal            440  65     70      60
```

```
## 60   54                   Psyduck   Water            320  50     52       48
## 61   55                   Golduck   Water            500  80     82       78
## 64   58                 Growlithe    Fire            350  55     70       45
## 65   59                  Arcanine    Fire            555  90    110       80
## 66   60                   Poliwag   Water            300  40     50       40
## 67   61                  Poliwhirl  Water            385  65     65       65
## 68   62                  Poliwrath  Water Fighting   510  90     95       95
## 69   63                       Abra Psychic           310  25     20       15
## 70   64                   Kadabra Psychic            400  40     35       30
## 71   65                   Alakazam Psychic           500  55     50       45
## 72   65     AlakazamMega Alakazam Psychic            590  55     50       65
## 76   69                 Bellsprout   Grass   Poison  300  50     75       35
## 77   70                 Weepinbell   Grass   Poison  390  65     90       50
## 78   71                 Victreebel   Grass   Poison  490  80    105       65
## 79   72                  Tentacool   Water   Poison  335  40     40       35
## 80   73                 Tentacruel   Water   Poison  515  80     70       65
## 84   77                     Ponyta    Fire            410  50     85       55
## 85   78                   Rapidash    Fire            500  65    100       70
## 86   79                   Slowpoke   Water  Psychic  315  90     65       65
## 87   80                    Slowbro   Water  Psychic  490  95     75      110
## 88   80       SlowbroMega Slowbro   Water  Psychic  590  95     75      180
## 91   83                 Farfetch'd  Normal   Flying  352  52     65       55
## 92   84                      Doduo  Normal   Flying  310  35     85       45
## 93   85                     Dodrio  Normal   Flying  460  60    110       70
## 94   86                       Seel   Water            325  65     45       55
## 95   87                    Dewgong   Water      Ice   475  90     70       80
## 98   90                   Shellder   Water            305  30     65      100
## 99   91                   Cloyster   Water      Ice   525  50     95      180
## 105  96                    Drowzee Psychic           328  60     48       45
## 106  97                       Hypno Psychic          483  85     73       70
## 107  98                     Krabby   Water            325  30    105       90
## 108  99                     Kingler   Water            475  55    130      115
## 111 102                   Exeggcute   Grass  Psychic  325  60     40       80
## 112 103                   Exeggutor   Grass  Psychic  520  95     95       85
## 117 108                   Lickitung  Normal           385  90     55       75
## 122 113                     Chansey  Normal           450 250      5        5
## 123 114                     Tangela   Grass           435  65     55      115
## 124 115                   Kangaskhan  Normal           490 105     95       80
## 125 115 KangaskhanMega Kangaskhan  Normal           590 105    125      100
## 126 116                      Horsea   Water            295  30     40       70
## 127 117                      Seadra   Water            440  55     65       95
## 128 118                     Goldeen   Water            320  45     67       60
## 129 119                     Seaking   Water            450  80     92       65
## 130 120                      Staryu   Water            340  30     45       55
## 131 121                     Starmie   Water  Psychic  520  60     75       85
## 132 122                   Mr. Mime Psychic    Fairy   460  40     45       65
## 133 123                     Scyther     Bug   Flying  500  70    110       80
## 136 126                     Magmar    Fire            495  65     95       57
## 137 127                     Pinsir     Bug            500  65    125      100
## 138 127       PinsirMega Pinsir     Bug   Flying   600  65    155      120
## 139 128                     Tauros  Normal           490  75    100       95
## 140 129                   Magikarp   Water            200  20     10       55
## 141 130                   Gyarados   Water   Flying  540  95    125       79
## 142 130     GyaradosMega Gyarados   Water     Dark  640  95    155      109
```

```
## 143 131                   Lapras   Water      Ice   535 130     85       80
## 144 132                   Ditto  Normal            288  48     48       48
## 145 133                   Eevee  Normal            325  55     55       50
## 146 134                 Vaporeon   Water            525 130     65       60
## 148 136                  Flareon    Fire            525  65    130       60
## 149 137                  Porygon  Normal            395  65     60       70
## 156 143                  Snorlax  Normal            540 160    110       65
## 159 146                  Moltres    Fire   Flying   580  90    100       90
## 163 150                   Mewtwo Psychic            680 106    110       90
## 164 150       MewtwoMega Mewtwo X Psychic Fighting   780 106    190      100
## 165 150       MewtwoMega Mewtwo Y Psychic            780 106    150       70
## 166 151                      Mew Psychic            600 100    100      100
## 167 152                Chikorita   Grass            318  45     49       65
## 168 153                  Bayleef   Grass            405  60     62       80
## 169 154                 Meganium   Grass            525  80     82      100
## 170 155                Cyndaquil    Fire            309  39     52       43
## 171 156                  Quilava    Fire            405  58     64       58
## 172 157                Typhlosion    Fire            534  78     84       78
## 173 158                  Totodile   Water            314  50     65       64
## 174 159                  Croconaw   Water            405  65     80       80
## 175 160                Feraligatr   Water            530  85    105      100
## 176 161                   Sentret  Normal            215  35     46       34
## 177 162                    Furret  Normal            415  85     76       64
## 178 163                  Hoothoot  Normal   Flying   262  60     30       30
## 179 164                  Noctowl  Normal   Flying   442 100     50       50
## 180 165                    Ledyba     Bug   Flying   265  40     20       30
## 181 166                    Ledian     Bug   Flying   390  55     35       50
## 182 167                  Spinarak     Bug   Poison   250  40     60       40
## 183 168                   Ariados     Bug   Poison   390  70     90       70
## 185 170                  Chinchou   Water Electric   330  75     38       38
## 186 171                   Lanturn   Water Electric   460 125     58       58
## 189 174                 Igglybuff  Normal    Fairy   210  90     30       15
## 192 177                      Natu Psychic   Flying   320  40     50       45
## 193 178                      Xatu Psychic   Flying   470  65     75       70
## 198 182                 Bellossom   Grass            490  75     80       95
## 199 183                    Marill   Water    Fairy   250  70     20       50
## 200 184                 Azumarill   Water    Fairy   420 100     50       80
## 202 186                  Politoed   Water            500  90     75       75
## 203 187                    Hoppip   Grass   Flying   250  35     35       40
## 204 188                  Skiploom   Grass   Flying   340  55     45       50
## 205 189                  Jumpluff   Grass   Flying   460  75     55       70
## 206 190                     Aipom  Normal            360  55     70       55
## 207 191                   Sunkern   Grass            180  30     30       30
## 208 192                  Sunflora   Grass            425  75     75       55
## 209 193                     Yanma     Bug   Flying   390  65     65       45
## 210 194                    Wooper   Water   Ground   210  55     45       45
## 211 195                  Quagsire   Water   Ground   430  95     85       85
## 212 196                    Espeon Psychic            525  65     65       60
## 215 199                  Slowking   Water  Psychic   490  95     75       80
## 217 201                     Unown Psychic            336  48     72       48
## 218 202                 Wobbuffet Psychic            405 190     33       58
## 219 203                 Girafarig  Normal  Psychic   455  70     80       65
## 220 204                    Pineco     Bug            290  50     65       90
## 221 205                Forretress     Bug    Steel   465  75     90      140
```

```
## 222 206                Dunsparce  Normal          415 100     70      70
## 228 211                 Qwilfish   Water   Poison 430  65     95      75
## 229 212                   Scizor     Bug    Steel 500  70    130     100
## 230 212        ScizorMega Scizor     Bug    Steel 600  70    150     140
## 231 213                  Shuckle     Bug     Rock 505  20     10     230
## 232 214                Heracross     Bug Fighting 500  80    125      75
## 233 214   HeracrossMega Heracross     Bug Fighting 600  80    185     115
## 235 216                Teddiursa  Normal          330  60     80      50
## 236 217                 Ursaring  Normal          500  90    130      75
## 237 218                   Slugma    Fire          250  40     40      40
## 238 219                 Magcargo    Fire     Rock 410  50     50     120
## 241 222                  Corsola   Water     Rock 380  55     55      85
## 242 223                 Remoraid   Water          300  35     65      35
## 243 224                Octillery   Water          480  75    105      75
## 245 226                  Mantine   Water   Flying 465  65     40      70
## 250 230                  Kingdra   Water   Dragon 540  75     95      95
## 253 233                 Porygon2  Normal          515  85     80      90
## 254 234                 Stantler  Normal          465  73     95      62
## 255 235                 Smeargle  Normal          250  55     20      35
## 260 240                    Magby    Fire          365  45     75      37
## 261 241                  Miltank  Normal          490  95     80     105
## 262 242                  Blissey  Normal          540 255     10      10
## 264 244                    Entei    Fire          580 115    115      85
## 265 245                  Suicune   Water          580 100     75     115
## 270 249                    Lugia Psychic   Flying 680 106     90     130
## 271 250                    Ho-oh    Fire   Flying 680 106    130      90
## 272 251                   Celebi Psychic    Grass 600 100    100     100
## 273 252                  Treecko   Grass          310  40     45      35
## 274 253                  Grovyle   Grass          405  50     65      45
## 275 254                 Sceptile   Grass          530  70     85      65
## 276 254    SceptileMega Sceptile   Grass   Dragon 630  70    110      75
## 277 255                  Torchic    Fire          310  45     60      40
## 278 256                Combusken    Fire Fighting 405  60     85      60
## 279 257                 Blaziken    Fire Fighting 530  80    120      70
## 280 257    BlazikenMega Blaziken    Fire Fighting 630  80    160      80
## 281 258                   Mudkip   Water          310  50     70      50
## 282 259                Marshtomp   Water   Ground 405  70     85      70
## 283 260                 Swampert   Water   Ground 535 100    110      90
## 284 260    SwampertMega Swampert   Water   Ground 635 100    150     110
## 287 263                Zigzagoon  Normal          240  38     30      41
## 288 264                  Linoone  Normal          420  78     70      61
## 289 265                  Wurmple     Bug          195  45     45      35
## 290 266                  Silcoon     Bug          205  50     35      55
## 291 267                Beautifly     Bug   Flying 395  60     70      50
## 292 268                  Cascoon     Bug          205  50     35      55
## 293 269                   Dustox     Bug   Poison 385  60     50      70
## 294 270                    Lotad   Water    Grass 220  40     30      30
## 295 271                   Lombre   Water    Grass 340  60     50      50
## 296 272                  Ludicolo   Water    Grass 480  80     70      70
## 297 273                   Seedot   Grass          220  40     40      50
## 298 274                  Nuzleaf   Grass     Dark 340  70     70      40
## 299 275                  Shiftry   Grass     Dark 480  90    100      60
## 300 276                  Taillow  Normal   Flying 270  40     55      30
## 301 277                  Swellow  Normal   Flying 430  60     85      60
```

```
## 302 278                  Wingull   Water   Flying   270  40     30      30
## 303 279                 Pelipper   Water   Flying   430  60     50     100
## 304 280                    Ralts Psychic    Fairy   198  28     25      25
## 305 281                   Kirlia Psychic    Fairy   278  38     35      35
## 306 282                 Gardevoir Psychic    Fairy   518  68     65      65
## 307 282   GardevoirMega Gardevoir Psychic    Fairy   618  68     85      65
## 308 283                  Surskit     Bug    Water   269  40     30      32
## 309 284                Masquerain     Bug   Flying   414  70     60      62
## 310 285                Shroomish   Grass            295  60     40      60
## 311 286                  Breloom   Grass Fighting   460  60    130      80
## 312 287                  Slakoth  Normal            280  60     60      60
## 313 288                 Vigoroth  Normal            440  80     80      80
## 314 289                  Slaking  Normal            670 150    160     100
## 315 290                  Nincada     Bug   Ground   266  31     45      90
## 316 291                  Ninjask     Bug   Flying   456  61     90      45
## 317 292                 Shedinja     Bug    Ghost   236   1     90      45
## 318 293                  Whismur  Normal            240  64     51      23
## 319 294                  Loudred  Normal            360  84     71      43
## 320 295                  Exploud  Normal            490 104     91      63
## 323 298                  Azurill  Normal    Fairy   190  50     20      40
## 325 300                   Skitty  Normal            260  50     45      45
## 326 301                 Delcatty  Normal            380  70     65      65
## 343 313                  Volbeat     Bug            400  65     73      55
## 344 314                 Illumise     Bug            400  65     47      55
## 345 315                  Roselia   Grass   Poison   400  50     60      45
## 348 318                 Carvanha   Water     Dark   305  45     90      20
## 349 319                 Sharpedo   Water     Dark   460  70    120      40
## 350 319    SharpedoMega Sharpedo   Water     Dark   560  70    140      70
## 351 320                  Wailmer   Water            400 130     70      35
## 352 321                  Wailord   Water            500 170     90      45
## 353 322                    Numel    Fire   Ground   305  60     60      40
## 354 323                 Camerupt    Fire   Ground   460  70    100      70
## 355 323    CameruptMega Camerupt    Fire   Ground   560  70    120     100
## 356 324                  Torkoal    Fire            470  70     85     140
## 357 325                   Spoink Psychic            330  60     25      35
## 358 326                  Grumpig Psychic            470  80     45      65
## 359 327                   Spinda  Normal            360  60     60      60
## 363 331                   Cacnea   Grass            335  50     85      40
## 364 332                 Cacturne   Grass     Dark   475  70    115      60
## 365 333                   Swablu  Normal   Flying   310  45     40      60
## 368 335                 Zangoose  Normal            458  73    115      60
## 372 339                 Barboach   Water   Ground   288  50     48      43
## 373 340                 Whiscash   Water   Ground   468 110     78      73
## 374 341                 Corphish   Water            308  43     80      65
## 375 342                Crawdaunt   Water     Dark   468  63    120      85
## 382 349                   Feebas   Water            200  20     15      20
## 383 350                  Milotic   Water            540  95     60      79
## 384 351                 Castform  Normal            420  70     70      70
## 385 352                  Kecleon  Normal            440  60     90      70
## 391 357                  Tropius   Grass   Flying   460  99     68      83
## 392 358                 Chimecho Psychic            425  65     50      70
## 395 360                   Wynaut Psychic            260  95     23      48
## 402 366                 Clamperl   Water            345  35     64      85
## 403 367                  Huntail   Water            485  55    104     105
```

```
## 404 368                 Gorebyss   Water            485  55     84     105
## 405 369                Relicanth   Water     Rock   485 100     90     130
## 406 370                  Luvdisc   Water            330  43     30      55
## 422 382                   Kyogre   Water            670 100    100      90
## 423 382       KyogrePrimal Kyogre   Water            770 100    150      90
## 429 386        DeoxysNormal Forme Psychic            600  50    150      50
## 430 386        DeoxysAttack Forme Psychic            600  50    180      20
## 431 386       DeoxysDefense Forme Psychic            600  50     70     160
## 432 386         DeoxysSpeed Forme Psychic            600  50     95      90
## 433 387                  Turtwig   Grass            318  55     68      64
## 434 388                   Grotle   Grass            405  75     89      85
## 435 389                 Torterra   Grass   Ground   525  95    109     105
## 436 390                 Chimchar    Fire            309  44     58      44
## 437 391                 Monferno    Fire Fighting   405  64     78      52
## 438 392                Infernape    Fire Fighting   534  76    104      71
## 439 393                   Piplup   Water            314  53     51      53
## 440 394                 Prinplup   Water            405  64     66      68
## 441 395                 Empoleon   Water    Steel   530  84     86      88
## 442 396                   Starly  Normal   Flying   245  40     55      30
## 443 397                 Staravia  Normal   Flying   340  55     75      50
## 444 398                Staraptor  Normal   Flying   485  85    120      70
## 445 399                   Bidoof  Normal            250  59     45      40
## 446 400                  Bibarel  Normal    Water   410  79     85      60
## 447 401                Kricketot     Bug            194  37     25      41
## 448 402                Kricketune     Bug            384  77     85      51
## 452 406                    Budew   Grass   Poison   280  40     30      35
## 453 407                  Roserade   Grass   Poison   515  60     70      65
## 458 412                    Burmy     Bug            224  40     29      45
## 459 413       WormadamPlant Cloak     Bug    Grass   424  60     59      85
## 460 413       WormadamSandy Cloak     Bug   Ground   424  60     79     105
## 461 413       WormadamTrash Cloak     Bug    Steel   424  60     69      95
## 462 414                  Mothim     Bug   Flying   424  70     94      50
## 463 415                  Combee     Bug   Flying   244  30     30      42
## 464 416                Vespiquen     Bug   Flying   474  70     80     102
## 466 418                   Buizel   Water            330  55     65      35
## 467 419                 Floatzel   Water            495  85    105      55
## 468 420                  Cherubi   Grass            275  45     35      45
## 469 421                  Cherrim   Grass            450  70     60      70
## 470 422                  Shellos   Water            325  76     48      48
## 471 423                 Gastrodon   Water   Ground   475 111     83      68
## 472 424                  Ambipom  Normal            482  75    100      66
## 475 427                  Buneary  Normal            350  55     66      44
## 476 428                  Lopunny  Normal            480  65     76      84
## 477 428       LopunnyMega Lopunny  Normal Fighting   580  65    136      94
## 480 431                  Glameow  Normal            310  49     55      42
## 481 432                  Purugly  Normal            452  71     82      64
## 482 433                Chingling Psychic            285  45     30      50
## 488 439                 Mime Jr. Psychic    Fairy   310  20     25      45
## 489 440                  Happiny  Normal            220 100      5       5
## 490 441                   Chatot  Normal   Flying   411  76     65      45
## 496 446                 Munchlax  Normal            390 135     85      40
## 506 455                Carnivine   Grass            454  74    100      72
## 507 456                  Finneon   Water            330  49     49      56
## 508 457                 Lumineon   Water            460  69     69      76
```

```
## 509 458                  Mantyke   Water   Flying   345  45     20      50
## 510 459                   Snover   Grass      Ice   334  60     62      50
## 511 460                Abomasnow   Grass      Ice   494  90     92      75
## 512 460   AbomasnowMega Abomasnow   Grass      Ice   594  90    132     105
## 515 463                Lickilicky  Normal           515 110     85      95
## 517 465                Tangrowth   Grass           535 100    100     125
## 519 467                Magmortar    Fire           540  75     95      67
## 521 469                  Yanmega     Bug   Flying   515  86     76      86
## 522 470                  Leafeon   Grass           525  65    110     130
## 526 474                Porygon-Z  Normal           535  85     80      70
## 527 475                  Gallade Psychic Fighting   518  68    125      65
## 528 475       GalladeMega Gallade Psychic Fighting   618  68    165      95
## 538 480                     Uxie Psychic           580  75     75     130
## 539 481                  Mesprit Psychic           580  80    105     105
## 540 482                    Azelf Psychic           580  75    125      70
## 542 484                   Palkia   Water   Dragon   680  90    120     100
## 543 485                   Heatran    Fire    Steel   600  91     90     106
## 544 486                Regigigas  Normal           670 110    160     110
## 547 488                Cresselia Psychic           600 120     70     120
## 548 489                   Phione   Water           480  80     80      80
## 549 490                  Manaphy   Water           600 100    100     100
## 551 492         ShayminLand Forme   Grass           600 100    100     100
## 552 492          ShayminSky Forme   Grass   Flying   600 100    103      75
## 553 493                   Arceus  Normal           720 120    120     120
## 554 494                  Victini Psychic     Fire   600 100    100     100
## 555 495                    Snivy   Grass           308  45     45      55
## 556 496                   Servine   Grass           413  60     60      75
## 557 497                 Serperior   Grass           528  75     75      95
## 558 498                    Tepig    Fire           308  65     63      45
## 559 499                   Pignite    Fire Fighting   418  90     93      55
## 560 500                   Emboar    Fire Fighting   528 110    123      65
## 561 501                  Oshawott   Water           308  55     55      45
## 562 502                    Dewott   Water           413  75     75      60
## 563 503                  Samurott   Water           528  95    100      85
## 564 504                    Patrat  Normal           255  45     55      39
## 565 505                   Watchog  Normal           420  60     85      69
## 566 506                  Lillipup  Normal           275  45     60      45
## 567 507                   Herdier  Normal           370  65     80      65
## 568 508                 Stoutland  Normal           500  85    110      90
## 571 511                   Pansage   Grass           316  50     53      48
## 572 512                   Simisage   Grass           498  75     98      63
## 573 513                   Pansear    Fire           316  50     53      48
## 574 514                   Simisear    Fire           498  75     98      63
## 575 515                   Panpour   Water           316  50     53      48
## 576 516                   Simipour   Water           498  75     98      63
## 577 517                      Munna Psychic           292  76     25      45
## 578 518                  Musharna Psychic           487 116     55      85
## 579 519                    Pidove  Normal   Flying   264  50     55      50
## 580 520                 Tranquill  Normal   Flying   358  62     77      62
## 581 521                 Unfezant  Normal   Flying   488  80    115      80
## 587 527                    Woobat Psychic   Flying   313  55     45      43
## 588 528                   Swoobat Psychic   Flying   425  67     57      55
## 591 531                    Audino  Normal           445 103     60      86
## 592 531          AudinoMega Audino  Normal    Fairy   545 103     60     126
```

```
## 596 535                  Tympole   Water          294  50     50       40
## 597 536                 Palpitoad   Water   Ground 384  75     65       55
## 598 537                Seismitoad   Water   Ground 509 105     95       75
## 601 540                  Sewaddle     Bug    Grass 310  45     53       70
## 602 541                  Swadloon     Bug    Grass 380  55     63       90
## 603 542                  Leavanny     Bug    Grass 500  75    103       80
## 604 543                  Venipede     Bug   Poison 260  30     45       59
## 605 544                 Whirlipede     Bug   Poison 360  40     55       99
## 606 545                 Scolipede     Bug   Poison 485  60    100       89
## 607 546                  Cottonee   Grass    Fairy 280  40     27       60
## 608 547                 Whimsicott   Grass    Fairy 480  60     67       85
## 609 548                    Petilil   Grass          280  45     35       50
## 610 549                 Lilligant   Grass          480  70     60       75
## 611 550                  Basculin   Water          460  70     92       65
## 615 554                  Darumaka    Fire          315  70     90       45
## 616 555   DarmanitanStandard Mode    Fire          480 105    140       55
## 617 555        DarmanitanZen Mode    Fire  Psychic 540 105     30      105
## 618 556                  Maractus   Grass          461  75     86       67
## 619 557                  Dwebble     Bug     Rock 325  50     65       85
## 620 558                  Crustle     Bug     Rock 475  70     95      125
## 623 561                  Sigilyph Psychic   Flying 490  72     58       80
## 626 564                  Tirtouga   Water     Rock 355  54     78      103
## 627 565                 Carracosta   Water     Rock 495  74    108      133
## 634 572                  Minccino  Normal          300  55     50       40
## 635 573                  Cinccino  Normal          470  75     95       60
## 636 574                    Gothita Psychic          290  45     30       50
## 637 575                  Gothorita Psychic          390  60     45       70
## 638 576                 Gothitelle Psychic          490  70     55       95
## 639 577                    Solosis Psychic          290  45     30       40
## 640 578                    Duosion Psychic          370  65     40       50
## 641 579                  Reuniclus Psychic          490 110     65       75
## 642 580                  Ducklett   Water   Flying 305  62     44       50
## 643 581                     Swanna   Water   Flying 473  75     87       63
## 647 585                  Deerling  Normal    Grass 335  60     60       50
## 648 586                  Sawsbuck  Normal    Grass 475  80    100       70
## 650 588                 Karrablast     Bug          315  50     75       45
## 651 589                 Escavalier     Bug    Steel 495  70    135      105
## 652 590                    Foongus   Grass   Poison 294  69     55       45
## 653 591                  Amoonguss   Grass   Poison 464 114     85       70
## 654 592                  Frillish   Water    Ghost 335  55     40       50
## 655 593                  Jellicent   Water    Ghost 480 100     60       70
## 656 594                  Alomomola   Water          470 165     75       80
## 657 595                     Joltik     Bug Electric 319  50     47       50
## 658 596                 Galvantula     Bug Electric 472  70     77       60
## 659 597                  Ferroseed   Grass    Steel 305  44     50       91
## 660 598                 Ferrothorn   Grass    Steel 489  74     94      131
## 667 605                     Elgyem Psychic          335  55     55       55
## 668 606                   Beheeyem Psychic          485  75     75       75
## 678 616                    Shelmet     Bug          305  50     40       85
## 679 617                  Accelgor     Bug          495  80     70       40
## 688 626                 Bouffalant  Normal          490  95    110       95
## 689 627                    Rufflet  Normal   Flying 350  70     83       50
## 690 628                  Braviary  Normal   Flying 510 100    123       75
## 693 631                    Heatmor    Fire          484  85     97       66
```

```
## 694 632                   Durant     Bug    Steel   484  58    109      112
## 698 636                 Larvesta     Bug     Fire   360  55     85       55
## 699 637                Volcarona     Bug     Fire   550  85     60       65
## 702 640                 Virizion   Grass Fighting   580  91     90       72
## 714 647     KeldeoOrdinary Forme   Water Fighting   580  91     72       90
## 715 647     KeldeoResolute Forme   Water Fighting   580  91     72       90
## 716 648       MeloettaAria Forme  Normal  Psychic   600 100     77       77
## 717 648  MeloettaPirouette Forme  Normal Fighting   600 100    128       90
## 718 649                 Genesect     Bug    Steel   600  71    120       95
## 719 650                  Chespin   Grass            313  56     61       65
## 720 651                Quilladin   Grass            405  61     78       95
## 721 652               Chesnaught   Grass Fighting   530  88    107      122
## 722 653                 Fennekin    Fire            307  40     45       40
## 723 654                  Braixen    Fire            409  59     59       58
## 724 655                  Delphox    Fire  Psychic   534  75     69       72
## 725 656                  Froakie   Water            314  41     56       40
## 726 657                Frogadier   Water            405  54     63       52
## 727 658                  Greninja   Water     Dark   530  72     95       67
## 728 659                  Bunnelby  Normal            237  38     36       38
## 729 660                 Diggersby  Normal   Ground   423  85     56       77
## 730 661                Fletchling  Normal   Flying   278  45     50       43
## 731 662               Fletchinder    Fire   Flying   382  62     73       55
## 732 663                Talonflame    Fire   Flying   499  78     81       71
## 733 664                Scatterbug     Bug            200  38     35       40
## 734 665                    Spewpa     Bug            213  45     22       60
## 735 666                  Vivillon     Bug   Flying   411  80     52       50
## 736 667                    Litleo    Fire   Normal   369  62     50       58
## 737 668                    Pyroar    Fire   Normal   507  86     68       72
## 741 672                    Skiddo   Grass            350  66     65       48
## 742 673                    Gogoat   Grass            531 123    100       62
## 745 676                   Furfrou  Normal            472  75     80       60
## 746 677                    Espurr Psychic            355  62     48       54
## 747 678              MeowsticMale Psychic            466  74     48       76
## 748 678            MeowsticFemale Psychic            466  74     48       76
## 763 692                 Clauncher   Water            330  50     53       62
## 764 693                 Clawitzer   Water            500  71     73       88
## 798 720       HoopaHoopa Confined Psychic    Ghost   600  80    110       60
## 799 720        HoopaHoopa Unbound Psychic     Dark   680  80    160       60
## 800 721                 Volcanion    Fire    Water   600  80    110      120
##     sp_atk sp_def speed generation legendary
## 1       65     65    45          1     False
## 2       80     80    60          1     False
## 3      100    100    80          1     False
## 4      122    120    80          1     False
## 5       60     50    65          1     False
## 6       80     65    80          1     False
## 7      109     85   100          1     False
## 8      130     85   100          1     False
## 9      159    115   100          1     False
## 10      50     64    43          1     False
## 11      65     80    58          1     False
## 12      85    105    78          1     False
## 13     135    115    78          1     False
## 14      20     20    45          1     False
```

```
## 15      25     25    30          1     False
## 16      90     80    70          1     False
## 17      20     20    50          1     False
## 18      25     25    35          1     False
## 19      45     80    75          1     False
## 20      15     80   145          1     False
## 21      35     35    56          1     False
## 22      50     50    71          1     False
## 23      70     70   101          1     False
## 24     135     80   121          1     False
## 25      25     35    72          1     False
## 26      50     70    97          1     False
## 27      31     31    70          1     False
## 28      61     61   100          1     False
## 43      50     65    65          1     False
## 44      81    100   100          1     False
## 45      45     25    20          1     False
## 46      85     50    45          1     False
## 49      75     65    30          1     False
## 50      85     75    40          1     False
## 51     110     90    50          1     False
## 52      45     55    25          1     False
## 53      60     80    30          1     False
## 54      40     55    45          1     False
## 55      90     75    90          1     False
## 58      40     40    90          1     False
## 59      65     65   115          1     False
## 60      65     50    55          1     False
## 61      95     80    85          1     False
## 64      70     50    60          1     False
## 65     100     80    95          1     False
## 66      40     40    90          1     False
## 67      50     50    90          1     False
## 68      70     90    70          1     False
## 69     105     55    90          1     False
## 70     120     70   105          1     False
## 71     135     95   120          1     False
## 72     175     95   150          1     False
## 76      70     30    40          1     False
## 77      85     45    55          1     False
## 78     100     70    70          1     False
## 79      50    100    70          1     False
## 80      80    120   100          1     False
## 84      65     65    90          1     False
## 85      80     80   105          1     False
## 86      40     40    15          1     False
## 87     100     80    30          1     False
## 88     130     80    30          1     False
## 91      58     62    60          1     False
## 92      35     35    75          1     False
## 93      60     60   100          1     False
## 94      45     70    45          1     False
## 95      70     95    70          1     False
## 98      45     25    40          1     False
```

```
## 99      85     45    70          1     False
## 105     43     90    42          1     False
## 106     73    115    67          1     False
## 107     25     25    50          1     False
## 108     50     50    75          1     False
## 111     60     45    40          1     False
## 112    125     65    55          1     False
## 117     60     75    30          1     False
## 122     35    105    50          1     False
## 123    100     40    60          1     False
## 124     40     80    90          1     False
## 125     60    100   100          1     False
## 126     70     25    60          1     False
## 127     95     45    85          1     False
## 128     35     50    63          1     False
## 129     65     80    68          1     False
## 130     70     55    85          1     False
## 131    100     85   115          1     False
## 132    100    120    90          1     False
## 133     55     80   105          1     False
## 136    100     85    93          1     False
## 137     55     70    85          1     False
## 138     65     90   105          1     False
## 139     40     70   110          1     False
## 140     15     20    80          1     False
## 141     60    100    81          1     False
## 142     70    130    81          1     False
## 143     85     95    60          1     False
## 144     48     48    48          1     False
## 145     45     65    55          1     False
## 146    110     95    65          1     False
## 148     95    110    65          1     False
## 149     85     75    40          1     False
## 156     65    110    30          1     False
## 159    125     85    90          1      True
## 163    154     90   130          1      True
## 164    154    100   130          1      True
## 165    194    120   140          1      True
## 166    100    100   100          1     False
## 167     49     65    45          2     False
## 168     63     80    60          2     False
## 169     83    100    80          2     False
## 170     60     50    65          2     False
## 171     80     65    80          2     False
## 172    109     85   100          2     False
## 173     44     48    43          2     False
## 174     59     63    58          2     False
## 175     79     83    78          2     False
## 176     35     45    20          2     False
## 177     45     55    90          2     False
## 178     36     56    50          2     False
## 179     76     96    70          2     False
## 180     40     80    55          2     False
## 181     55    110    85          2     False
```

```
## 182     40     40     30          2     False
## 183     60     60     40          2     False
## 185     56     56     67          2     False
## 186     76     76     67          2     False
## 189     40     20     15          2     False
## 192     70     45     70          2     False
## 193     95     70     95          2     False
## 198     90    100     50          2     False
## 199     20     50     40          2     False
## 200     60     80     50          2     False
## 202     90    100     70          2     False
## 203     35     55     50          2     False
## 204     45     65     80          2     False
## 205     55     95    110          2     False
## 206     40     55     85          2     False
## 207     30     30     30          2     False
## 208    105     85     30          2     False
## 209     75     45     95          2     False
## 210     25     25     15          2     False
## 211     65     65     35          2     False
## 212    130     95    110          2     False
## 215    100    110     30          2     False
## 217     72     48     48          2     False
## 218     33     58     33          2     False
## 219     90     65     85          2     False
## 220     35     35     15          2     False
## 221     60     60     40          2     False
## 222     65     65     45          2     False
## 228     55     55     85          2     False
## 229     55     80     65          2     False
## 230     65    100     75          2     False
## 231     10    230      5          2     False
## 232     40     95     85          2     False
## 233     40    105     75          2     False
## 235     50     50     40          2     False
## 236     75     75     55          2     False
## 237     70     40     20          2     False
## 238     80     80     30          2     False
## 241     65     85     35          2     False
## 242     65     35     65          2     False
## 243    105     75     45          2     False
## 245     80    140     70          2     False
## 250     95     95     85          2     False
## 253    105     95     60          2     False
## 254     85     65     85          2     False
## 255     20     45     75          2     False
## 260     70     55     83          2     False
## 261     40     70    100          2     False
## 262     75    135     55          2     False
## 264     90     75    100          2      True
## 265     90    115     85          2      True
## 270     90    154    110          2      True
## 271    110    154     90          2      True
## 272    100    100    100          2     False
```

```
## 273     65     55     70          3     False
## 274     85     65     95          3     False
## 275    105     85    120          3     False
## 276    145     85    145          3     False
## 277     70     50     45          3     False
## 278     85     60     55          3     False
## 279    110     70     80          3     False
## 280    130     80    100          3     False
## 281     50     50     40          3     False
## 282     60     70     50          3     False
## 283     85     90     60          3     False
## 284     95    110     70          3     False
## 287     30     41     60          3     False
## 288     50     61    100          3     False
## 289     20     30     20          3     False
## 290     25     25     15          3     False
## 291    100     50     65          3     False
## 292     25     25     15          3     False
## 293     50     90     65          3     False
## 294     40     50     30          3     False
## 295     60     70     50          3     False
## 296     90    100     70          3     False
## 297     30     30     30          3     False
## 298     60     40     60          3     False
## 299     90     60     80          3     False
## 300     30     30     85          3     False
## 301     50     50    125          3     False
## 302     55     30     85          3     False
## 303     85     70     65          3     False
## 304     45     35     40          3     False
## 305     65     55     50          3     False
## 306    125    115     80          3     False
## 307    165    135    100          3     False
## 308     50     52     65          3     False
## 309     80     82     60          3     False
## 310     40     60     35          3     False
## 311     60     60     70          3     False
## 312     35     35     30          3     False
## 313     55     55     90          3     False
## 314     95     65    100          3     False
## 315     30     30     40          3     False
## 316     50     50    160          3     False
## 317     30     30     40          3     False
## 318     51     23     28          3     False
## 319     71     43     48          3     False
## 320     91     73     68          3     False
## 323     20     40     20          3     False
## 325     35     35     50          3     False
## 326     55     55     70          3     False
## 343     47     75     85          3     False
## 344     73     75     85          3     False
## 345    100     80     65          3     False
## 348     65     20     65          3     False
## 349     95     40     95          3     False
```

```
## 350    110     65   105          3     False
## 351     70     35    60          3     False
## 352     90     45    60          3     False
## 353     65     45    35          3     False
## 354    105     75    40          3     False
## 355    145    105    20          3     False
## 356     85     70    20          3     False
## 357     70     80    60          3     False
## 358     90    110    80          3     False
## 359     60     60    60          3     False
## 363     85     40    35          3     False
## 364    115     60    55          3     False
## 365     40     75    50          3     False
## 368     60     60    90          3     False
## 372     46     41    60          3     False
## 373     76     71    60          3     False
## 374     50     35    35          3     False
## 375     90     55    55          3     False
## 382     10     55    80          3     False
## 383    100    125    81          3     False
## 384     70     70    70          3     False
## 385     60    120    40          3     False
## 391     72     87    51          3     False
## 392     95     80    65          3     False
## 395     23     48    23          3     False
## 402     74     55    32          3     False
## 403     94     75    52          3     False
## 404    114     75    52          3     False
## 405     45     65    55          3     False
## 406     40     65    97          3     False
## 422    150    140    90          3      True
## 423    180    160    90          3      True
## 429    150     50   150          3      True
## 430    180     20   150          3      True
## 431     70    160    90          3      True
## 432     95     90   180          3      True
## 433     45     55    31          4     False
## 434     55     65    36          4     False
## 435     75     85    56          4     False
## 436     58     44    61          4     False
## 437     78     52    81          4     False
## 438    104     71   108          4     False
## 439     61     56    40          4     False
## 440     81     76    50          4     False
## 441    111    101    60          4     False
## 442     30     30    60          4     False
## 443     40     40    80          4     False
## 444     50     60   100          4     False
## 445     35     40    31          4     False
## 446     55     60    71          4     False
## 447     25     41    25          4     False
## 448     55     51    65          4     False
## 452     50     70    55          4     False
## 453    125    105    90          4     False
```

```
## 458     29     45     36          4     False
## 459     79    105     36          4     False
## 460     59     85     36          4     False
## 461     69     95     36          4     False
## 462     94     50     66          4     False
## 463     30     42     70          4     False
## 464     80    102     40          4     False
## 466     60     30     85          4     False
## 467     85     50    115          4     False
## 468     62     53     35          4     False
## 469     87     78     85          4     False
## 470     57     62     34          4     False
## 471     92     82     39          4     False
## 472     60     66    115          4     False
## 475     44     56     85          4     False
## 476     54     96    105          4     False
## 477     54     96    135          4     False
## 480     42     37     85          4     False
## 481     64     59    112          4     False
## 482     65     50     45          4     False
## 488     70     90     60          4     False
## 489     15     65     30          4     False
## 490     92     42     91          4     False
## 496     40     85      5          4     False
## 506     90     72     46          4     False
## 507     49     61     66          4     False
## 508     69     86     91          4     False
## 509     60    120     50          4     False
## 510     62     60     40          4     False
## 511     92     85     60          4     False
## 512    132    105     30          4     False
## 515     80     95     50          4     False
## 517    110     50     50          4     False
## 519    125     95     83          4     False
## 521    116     56     95          4     False
## 522     60     65     95          4     False
## 526    135     75     90          4     False
## 527     65    115     80          4     False
## 528     65    115    110          4     False
## 538     75    130     95          4      True
## 539    105    105     80          4      True
## 540    125     70    115          4      True
## 542    150    120    100          4      True
## 543    130    106     77          4      True
## 544     80    110    100          4      True
## 547     75    130     85          4     False
## 548     80     80     80          4     False
## 549    100    100    100          4     False
## 551    100    100    100          4      True
## 552    120     75    127          4      True
## 553    120    120    120          4      True
## 554    100    100    100          5      True
## 555     45     55     63          5     False
## 556     60     75     83          5     False
```

```
## 557     75     95   113          5     False
## 558     45     45    45          5     False
## 559     70     55    55          5     False
## 560    100     65    65          5     False
## 561     63     45    45          5     False
## 562     83     60    60          5     False
## 563    108     70    70          5     False
## 564     35     39    42          5     False
## 565     60     69    77          5     False
## 566     25     45    55          5     False
## 567     35     65    60          5     False
## 568     45     90    80          5     False
## 571     53     48    64          5     False
## 572     98     63   101          5     False
## 573     53     48    64          5     False
## 574     98     63   101          5     False
## 575     53     48    64          5     False
## 576     98     63   101          5     False
## 577     67     55    24          5     False
## 578    107     95    29          5     False
## 579     36     30    43          5     False
## 580     50     42    65          5     False
## 581     65     55    93          5     False
## 587     55     43    72          5     False
## 588     77     55   114          5     False
## 591     60     86    50          5     False
## 592     80    126    50          5     False
## 596     50     40    64          5     False
## 597     65     55    69          5     False
## 598     85     75    74          5     False
## 601     40     60    42          5     False
## 602     50     80    42          5     False
## 603     70     80    92          5     False
## 604     30     39    57          5     False
## 605     40     79    47          5     False
## 606     55     69   112          5     False
## 607     37     50    66          5     False
## 608     77     75   116          5     False
## 609     70     50    30          5     False
## 610    110     75    90          5     False
## 611     80     55    98          5     False
## 615     15     45    50          5     False
## 616     30     55    95          5     False
## 617    140    105    55          5     False
## 618    106     67    60          5     False
## 619     35     35    55          5     False
## 620     65     75    45          5     False
## 623    103     80    97          5     False
## 626     53     45    22          5     False
## 627     83     65    32          5     False
## 634     40     40    75          5     False
## 635     65     60   115          5     False
## 636     55     65    45          5     False
## 637     75     85    55          5     False
```

```
## 638     95    110    65          5     False
## 639    105     50    20          5     False
## 640    125     60    30          5     False
## 641    125     85    30          5     False
## 642     44     50    55          5     False
## 643     87     63    98          5     False
## 647     40     50    75          5     False
## 648     60     70    95          5     False
## 650     40     45    60          5     False
## 651     60    105    20          5     False
## 652     55     55    15          5     False
## 653     85     80    30          5     False
## 654     65     85    40          5     False
## 655     85    105    60          5     False
## 656     40     45    65          5     False
## 657     57     50    65          5     False
## 658     97     60   108          5     False
## 659     24     86    10          5     False
## 660     54    116    20          5     False
## 667     85     55    30          5     False
## 668    125     95    40          5     False
## 678     40     65    25          5     False
## 679    100     60   145          5     False
## 688     40     95    55          5     False
## 689     37     50    60          5     False
## 690     57     75    80          5     False
## 693    105     66    65          5     False
## 694     48     48   109          5     False
## 698     50     55    60          5     False
## 699    135    105   100          5     False
## 702     90    129   108          5      True
## 714    129     90   108          5     False
## 715    129     90   108          5     False
## 716    128    128    90          5     False
## 717     77     77   128          5     False
## 718    120     95    99          5     False
## 719     48     45    38          6     False
## 720     56     58    57          6     False
## 721     74     75    64          6     False
## 722     62     60    60          6     False
## 723     90     70    73          6     False
## 724    114    100   104          6     False
## 725     62     44    71          6     False
## 726     83     56    97          6     False
## 727    103     71   122          6     False
## 728     32     36    57          6     False
## 729     50     77    78          6     False
## 730     40     38    62          6     False
## 731     56     52    84          6     False
## 732     74     69   126          6     False
## 733     27     25    35          6     False
## 734     27     30    29          6     False
## 735     90     50    89          6     False
## 736     73     54    72          6     False
```

```
## 737    109     66   106          6     False
## 741     62     57    52          6     False
## 742     97     81    68          6     False
## 745     65     90   102          6     False
## 746     63     60    68          6     False
## 747     83     81   104          6     False
## 748     83     81   104          6     False
## 763     58     63    44          6     False
## 764    120     89    59          6     False
## 798    150    130    70          6      True
## 799    170    130    80          6      True
## 800    130     90    70          6      True
```

```
names <- c('type_1' ,'legendary', 'generation')
pokemon_filter[,names] <- lapply(pokemon_filter[,names] , factor)
str(pokemon_filter)
```

```
## 'data.frame':    458 obs. of  13 variables:
##  $ x         : int  1 2 3 3 4 5 6 6 6 7 ...
##  $ name      : chr  "Bulbasaur" "Ivysaur" "Venusaur" "VenusaurMega Venusaur" ...
##  $ type_1    : Factor w/ 6 levels "Bug","Fire","Grass",..: 3 3 3 3 2 2 2 2 2 6 ...
##  $ type_2    : chr  "Poison" "Poison" "Poison" "Poison" ...
##  $ total     : int  318 405 525 625 309 405 534 634 634 314 ...
##  $ hp        : int  45 60 80 80 39 58 78 78 78 44 ...
##  $ attack    : int  49 62 82 100 52 64 84 130 104 48 ...
##  $ defense   : int  49 63 83 123 43 58 78 111 78 65 ...
##  $ sp_atk    : int  65 80 100 122 60 80 109 130 159 50 ...
##  $ sp_def    : int  65 80 100 120 50 65 85 85 115 64 ...
##  $ speed     : int  45 60 80 80 65 80 100 100 100 43 ...
##  $ generation: Factor w/ 6 levels "1","2","3","4",..: 1 1 1 1 1 1 1 1 1 1 ...
##  $ legendary : Factor w/ 2 levels "False","True": 1 1 1 1 1 1 1 1 1 1 ...
```

```
print(str(pokemon_filter))
```

```
## 'data.frame':    458 obs. of  13 variables:
##  $ x         : int  1 2 3 3 4 5 6 6 6 7 ...
##  $ name      : chr  "Bulbasaur" "Ivysaur" "Venusaur" "VenusaurMega Venusaur" ...
##  $ type_1    : Factor w/ 6 levels "Bug","Fire","Grass",..: 3 3 3 3 2 2 2 2 2 6 ...
##  $ type_2    : chr  "Poison" "Poison" "Poison" "Poison" ...
##  $ total     : int  318 405 525 625 309 405 534 634 634 314 ...
##  $ hp        : int  45 60 80 80 39 58 78 78 78 44 ...
##  $ attack    : int  49 62 82 100 52 64 84 130 104 48 ...
##  $ defense   : int  49 63 83 123 43 58 78 111 78 65 ...
##  $ sp_atk    : int  65 80 100 122 60 80 109 130 159 50 ...
##  $ sp_def    : int  65 80 100 120 50 65 85 85 115 64 ...
##  $ speed     : int  45 60 80 80 65 80 100 100 100 43 ...
##  $ generation: Factor w/ 6 levels "1","2","3","4",..: 1 1 1 1 1 1 1 1 1 1 ...
##  $ legendary : Factor w/ 2 levels "False","True": 1 1 1 1 1 1 1 1 1 1 ...
## NULL
```

```
set.seed(3435)
pokemon_split <- initial_split(pokemon_filter, strata = "type_1")
```

```
pokemon_train <- training(pokemon_split)
pokemon_test <- testing(pokemon_split)

pokemon_fold <- vfold_cv(pokemon_train, v = 5, strata = "type_1")
pokemon_fold
```

```
## #  5-fold cross-validation using stratification
## # A tibble: 5 x 2
##   splits           id
##   <list>           <chr>
## 1 <split [270/71]> Fold1
## 2 <split [271/70]> Fold2
## 3 <split [273/68]> Fold3
## 4 <split [274/67]> Fold4
## 5 <split [276/65]> Fold5
```

```
pokemon_recipe <- recipe(type_1 ~ legendary + generation + sp_atk + attack + speed + defense + hp + sp_
  step_dummy(all_nominal_predictors()) %>%
  step_normalize(all_numeric())
```

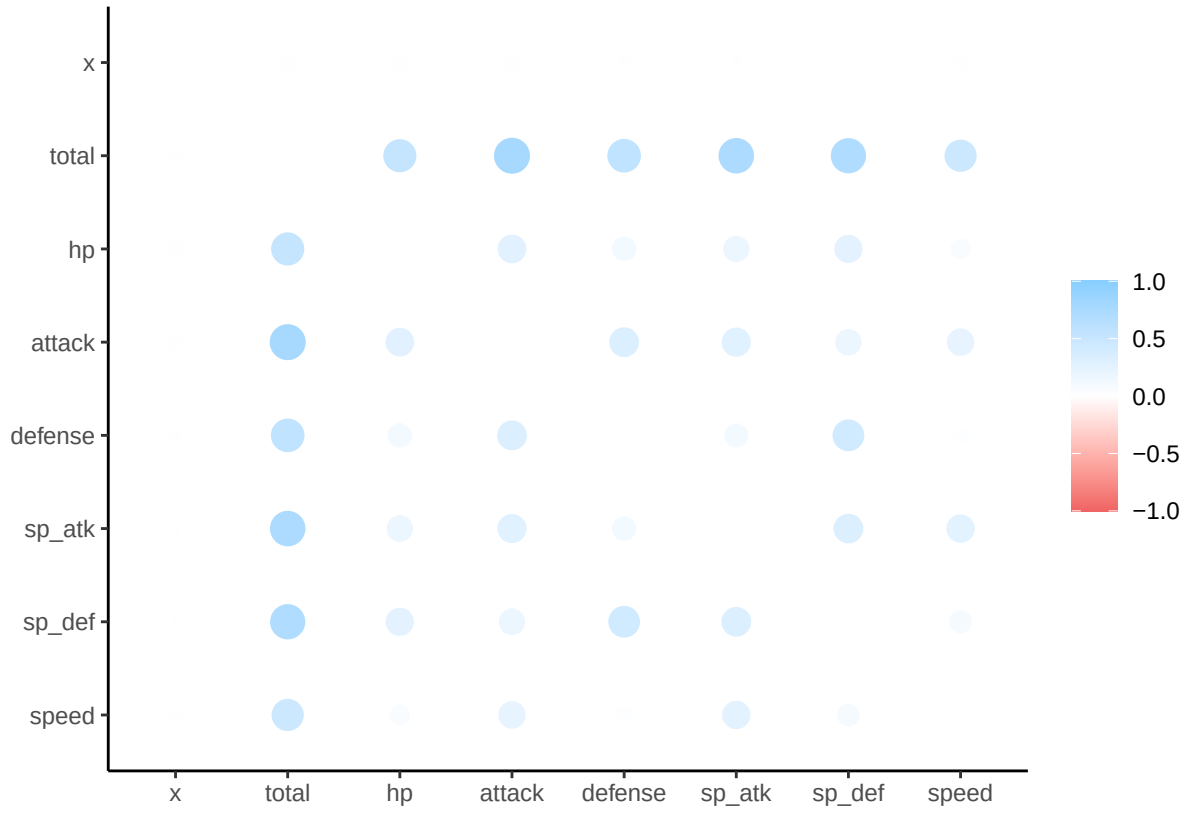
Exercise 2

Create a correlation matrix of the training set, using the `corrplot` package. *Note: You can choose how to handle the continuous variables for this plot; justify your decision(s).*

What relationships, if any, do you notice? Do these relationships make sense to you?

We can see that there is no negative correlation between the variables. Additionally, each variable has the strongest correlation with the 'total' variable. These relationships do make sense for me.

```
cor_pokemon_train <- pokemon_train %>%
  select(is.numeric) %>%
  cor(use = "pairwise.complete.obs", method = "pearson")
rplot(cor_pokemon_train)
```



Exercise 3

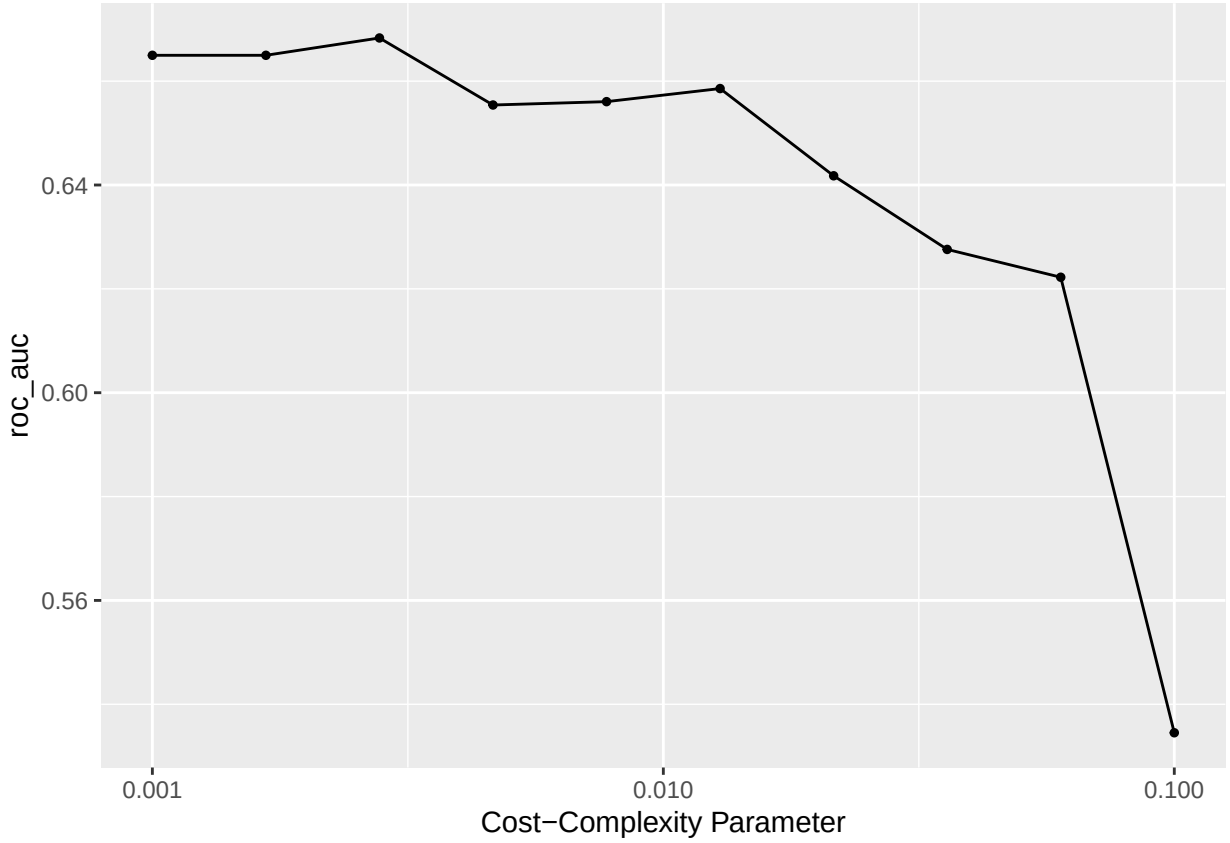
First, set up a decision tree model and workflow. Tune the `cost_complexity` hyperparameter. Use the same levels we used in Lab 7 – that is, `range = c(-3, -1)`. Specify that the metric we want to optimize is `roc_auc`.

Print an `autoplot()` of the results. What do you observe? Does a single decision tree perform better with a smaller or larger complexity penalty?

A single decision tree perform better with a smaller complexity penality because when the values are smaller, the 'roc_auc' values are larger. Larger 'roc_auc' values brings a better decision tree.

```
tree_spec <- decision_tree() %>%
  set_engine("rpart")

class_tree_spec <- tree_spec %>%
  set_mode("classification")

tree_workflow <- workflow() %>%
  add_model(class_tree_spec %>% set_args(cost_complexity = tune())) %>%
  add_recipe(pokemon_recipe)

set.seed(3435)
pokemon_fold <- vfold_cv(pokemon_train)

param_grid <- grid_regular(cost_complexity(range = c(-3, -1)), levels = 10)
```

```
tune_res <- tune_grid(
  tree_workflow,
  resamples = pokemon_fold,
  grid = param_grid,
  metrics = metric_set(roc_auc)
)

autoplot(tune_res)
```



Exercise 4

What is the `roc_auc` of your best-performing pruned decision tree on the folds? *Hint: Use `collect_metrics()` and `arrange()`.*

The 'roc_auc' of my best-performing pruned decision tree on the folds would be 0.6683159.

```
collect_metrics(tune_res)
```

```
## # A tibble: 10 x 7
##    cost_complexity .metric .estimator  mean     n std_err .config
##              <dbl> <chr>   <chr>      <dbl> <int>   <dbl> <chr>
##  1         0.001   roc_auc hand_till  0.665    10  0.0114 Preprocessor1_Model01
##  2         0.00167 roc_auc hand_till  0.665    10  0.0114 Preprocessor1_Model02
##  3         0.00278 roc_auc hand_till  0.668    10  0.0109 Preprocessor1_Model03
```

```
##  4         0.00464 roc_auc hand_till  0.655    10  0.0133 Preprocessor1_Model04
##  5         0.00774 roc_auc hand_till  0.656    10  0.0135 Preprocessor1_Model05
##  6         0.0129  roc_auc hand_till  0.659    10  0.0182 Preprocessor1_Model06
##  7         0.0215  roc_auc hand_till  0.642    10  0.0167 Preprocessor1_Model07
##  8         0.0359  roc_auc hand_till  0.628    10  0.0163 Preprocessor1_Model08
##  9         0.0599  roc_auc hand_till  0.622    10  0.0153 Preprocessor1_Model09
## 10         0.1     roc_auc hand_till  0.535    10  0.0159 Preprocessor1_Model10
```

```
arrange(tune_res)
```

```
## # Tuning results
## # 10-fold cross-validation
## # A tibble: 10 x 4
##    splits           id     .metrics          .notes
##    <list>           <chr>  <list>            <list>
##  1 <split [306/35]> Fold01 <tibble [10 x 5]> <tibble [0 x 3]>
##  2 <split [307/34]> Fold02 <tibble [10 x 5]> <tibble [0 x 3]>
##  3 <split [307/34]> Fold03 <tibble [10 x 5]> <tibble [0 x 3]>
##  4 <split [307/34]> Fold04 <tibble [10 x 5]> <tibble [0 x 3]>
##  5 <split [307/34]> Fold05 <tibble [10 x 5]> <tibble [0 x 3]>
##  6 <split [307/34]> Fold06 <tibble [10 x 5]> <tibble [0 x 3]>
##  7 <split [307/34]> Fold07 <tibble [10 x 5]> <tibble [0 x 3]>
##  8 <split [307/34]> Fold08 <tibble [10 x 5]> <tibble [0 x 3]>
##  9 <split [307/34]> Fold09 <tibble [10 x 5]> <tibble [0 x 3]>
## 10 <split [307/34]> Fold10 <tibble [10 x 5]> <tibble [0 x 3]>
```

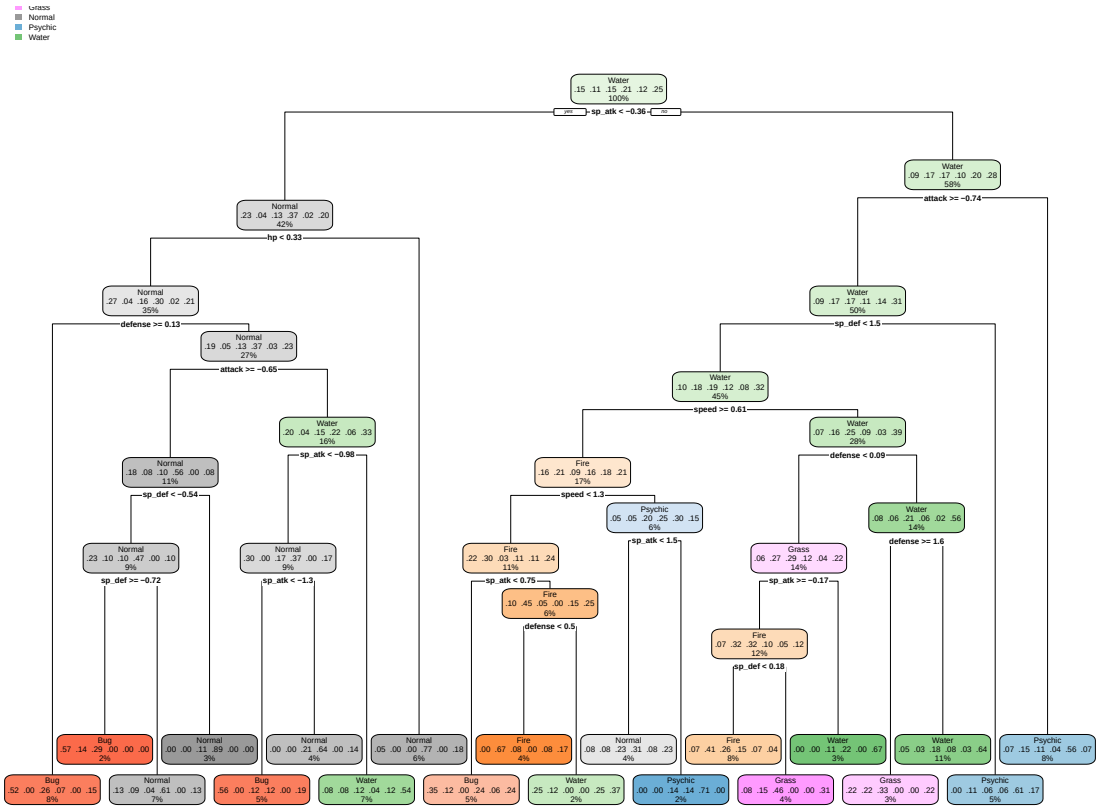
```
best_complexity <- select_best(tune_res)

best_complexity
```

```
## # A tibble: 1 x 2
##   cost_complexity .config
##             <dbl> <chr>
## 1         0.00278 Preprocessor1_Model03
```

Exercise 5

Using `rpart.plot`, fit and visualize your best-performing pruned decision tree with the *training* set.

```
class_tree_final <- finalize_workflow(tree_workflow, best_complexity)

class_tree_final_fit <- fit(class_tree_final, data = pokemon_train)

class_tree_final_fit %>%
  extract_fit_engine() %>%
  rpart.plot()
```



Exercise 5

Now set up a random forest model and workflow. Use the `ranger` engine and set `importance = "impurity"`. Tune `mtry`, `trees`, and `min_n`. Using the documentation for `rand_forest()`, explain in your own words what each of these hyperparameters represent.

The hyperparamter 'mtry' represent the number of variables randomly sampled as candidates at each split. 'trees' represent the number of trees to grow and 'min_n' represent the number of observations needed to keep splitting nodes.

Create a regular grid with 8 levels each. You can choose plausible ranges for each hyperparameter. Note that `mtry` should not be smaller than 1 or larger than 8. **Explain why not. What type of model would `mtry = 8` represent?**

'mtry = 8' would represent the creation of the tree and looking at 8 of my features before going into the next step.

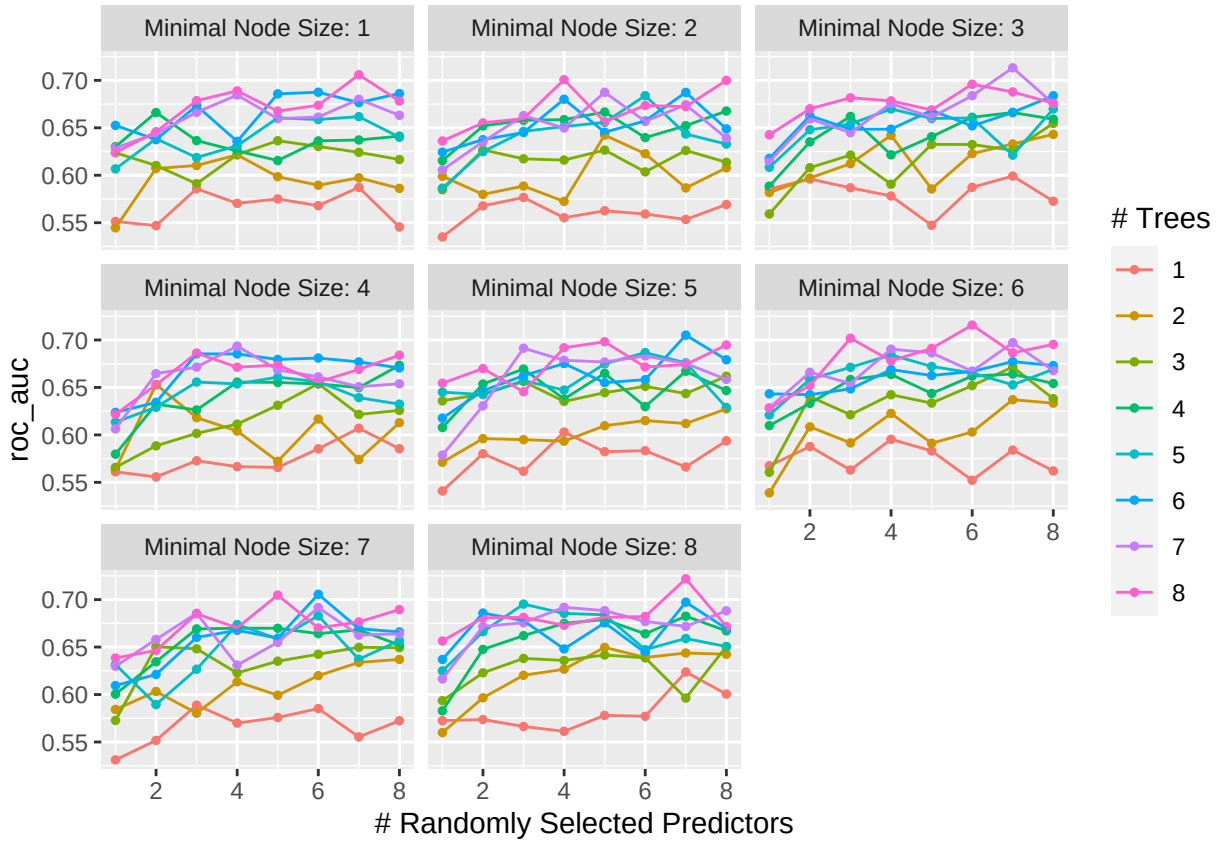
```
class_forest_spec <- rand_forest() %>%
  set_engine("ranger", importance = "impurity") %>%
  set_mode("classification")

param_grid2 <- grid_regular(mtry(range = c(1, 8)), trees(range = c(1,8)), min_n(range = c(1,8)),  levels

forest_workflow <- workflow() %>%
  add_model(class_forest_spec %>% set_args(mtry = tune(), trees = tune(), min_n = tune())) %>%
  add_recipe(pokemon_recipe)
```

```
####

#class_forest_spec <- rand_forest() %>%
#  set_engine("ranger", importance = "impurity") %>%
#  set_mode("classification")

#param_grid2_trees <- grid_regular(trees(range = c(1, 8)),  levels = 8)

#forest_workflow_trees <- workflow() %>%
#  add_model(class_forest_spec %>% set_args(trees = tune())) %>%
#  add_recipe(pokemon_recipe)

####
#class_forest_spec <- rand_forest() %>%
#  set_engine("ranger", importance = "impurity") %>%
#  set_mode("classification")

#param_grid2_min_n <- grid_regular(min_n(range = c(1, 8)),  levels = 8)

#forest_workflow_min_n <- workflow() %>%
#  add_model(class_forest_spec %>% set_args(min_n = tune())) %>%
#  add_recipe(pokemon_recipe)
```

Exercise 6

Specify `roc_auc` as a metric. Tune the model and print an `autoplot()` of the results. What do you observe? What values of the hyperparameters seem to yield the best performance?

```
tune_res_forest <- tune_grid(
  forest_workflow,
  resamples = pokemon_fold,
  grid = param_grid2,
  metrics = metric_set(roc_auc)
)

autoplot(tune_res_forest)
```




```
# ###
#
# tune_res_forest <- tune_grid(
#   forest_workflow_trees,
#   resamples = pokemon_fold,
#   grid = param_grid2_trees,
#   metrics = metric_set(roc_auc)
# )
#
# autoplot(tune_res_forest)
#
# ####
#
# tune_res_forest <- tune_grid(
#   forest_workflow_min_n,
#   resamples = pokemon_fold,
#   grid = param_grid2_min_n,
#   metrics = metric_set(roc_auc)
# )
#
# autoplot(tune_res_forest)
```

Exercise 7

What is the `roc_auc` of your best-performing random forest model on the folds? *Hint: Use `collect_metrics()` and `arrange()`.*

The 'roc_auc' of the best-performing random forest model on the folds would be 0.7219057.

```
collect_metrics(tune_res_forest)
```

```
## # A tibble: 512 x 9
##     mtry trees min_n .metric .estimator  mean     n std_err .config
##    <int> <int> <int> <chr>   <chr>      <dbl> <int>   <dbl> <chr>
##  1     1     1     1 roc_auc hand_till  0.551    10  0.0285 Preprocessor1_Model~
##  2     2     1     1 roc_auc hand_till  0.547    10  0.0248 Preprocessor1_Model~
##  3     3     1     1 roc_auc hand_till  0.586    10  0.0220 Preprocessor1_Model~
##  4     4     1     1 roc_auc hand_till  0.570    10  0.0154 Preprocessor1_Model~
##  5     5     1     1 roc_auc hand_till  0.575    10  0.0150 Preprocessor1_Model~
##  6     6     1     1 roc_auc hand_till  0.568    10  0.0155 Preprocessor1_Model~
##  7     7     1     1 roc_auc hand_till  0.587    10  0.0213 Preprocessor1_Model~
##  8     8     1     1 roc_auc hand_till  0.546    10  0.0170 Preprocessor1_Model~
##  9     1     2     1 roc_auc hand_till  0.545    10  0.0249 Preprocessor1_Model~
## 10     2     2     1 roc_auc hand_till  0.607    10  0.0197 Preprocessor1_Model~
## # ... with 502 more rows
```

```
arrange(tune_res_forest)
```

```
## # Tuning results
## # 10-fold cross-validation
## # A tibble: 10 x 4
##    splits           id     .metrics           .notes
##    <list>           <chr>  <list>             <list>
##  1 <split [306/35]> Fold01 <tibble [512 x 7]> <tibble [0 x 3]>
##  2 <split [307/34]> Fold02 <tibble [512 x 7]> <tibble [0 x 3]>
##  3 <split [307/34]> Fold03 <tibble [512 x 7]> <tibble [0 x 3]>
##  4 <split [307/34]> Fold04 <tibble [512 x 7]> <tibble [0 x 3]>
##  5 <split [307/34]> Fold05 <tibble [512 x 7]> <tibble [0 x 3]>
##  6 <split [307/34]> Fold06 <tibble [512 x 7]> <tibble [0 x 3]>
##  7 <split [307/34]> Fold07 <tibble [512 x 7]> <tibble [0 x 3]>
##  8 <split [307/34]> Fold08 <tibble [512 x 7]> <tibble [0 x 3]>
##  9 <split [307/34]> Fold09 <tibble [512 x 7]> <tibble [0 x 3]>
## 10 <split [307/34]> Fold10 <tibble [512 x 7]> <tibble [0 x 3]>
```

```
best_complexity2 <- select_best(tune_res_forest)
best_complexity2
```

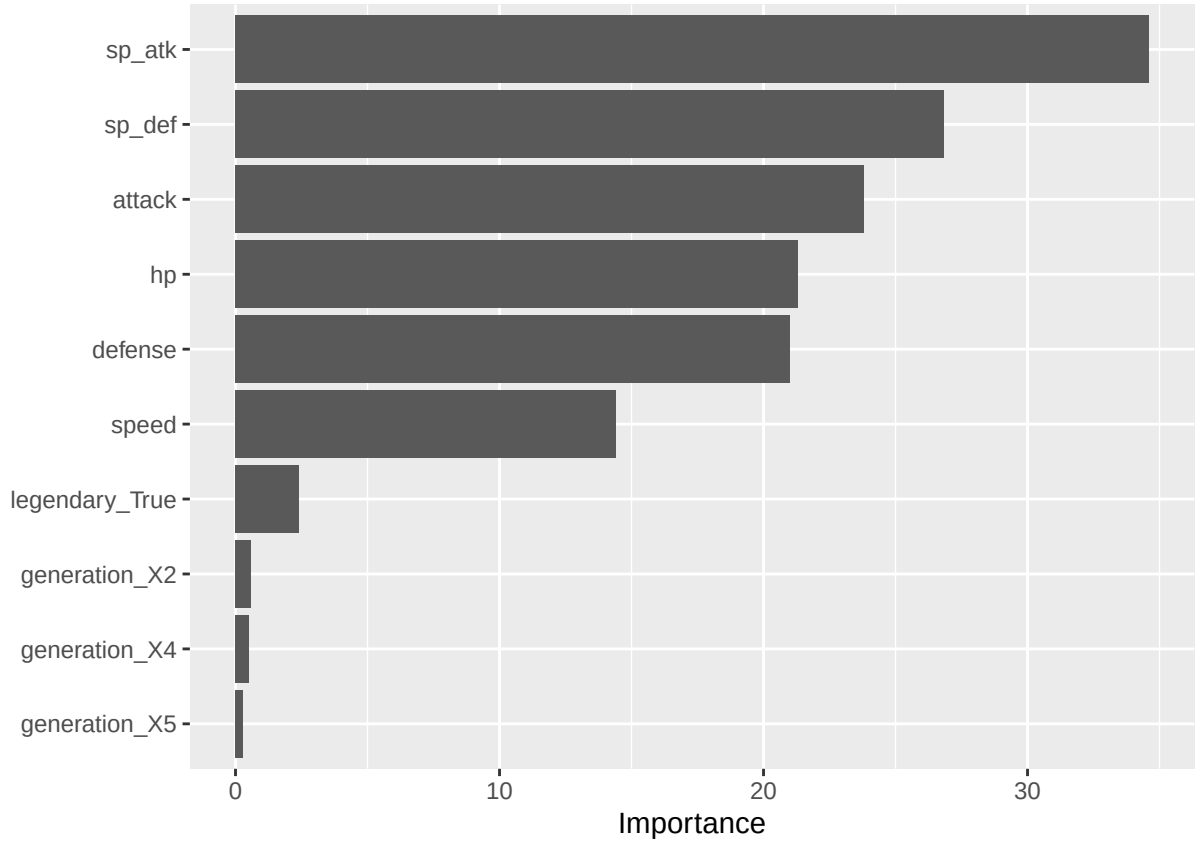
```
## # A tibble: 1 x 4
##    mtry trees min_n .config
##   <int> <int> <int> <chr>
## 1     7     8     8 Preprocessor1_Model511
```

Exercise 8

Create a variable importance plot, using `vip()`, with your best-performing random forest model fit on the *training* set.

Which variables were most useful? Which were least useful? Are these results what you expected, or not?

```
class_tree_final_fit %>%
  pull_workflow_fit() %>%
  vip()
```



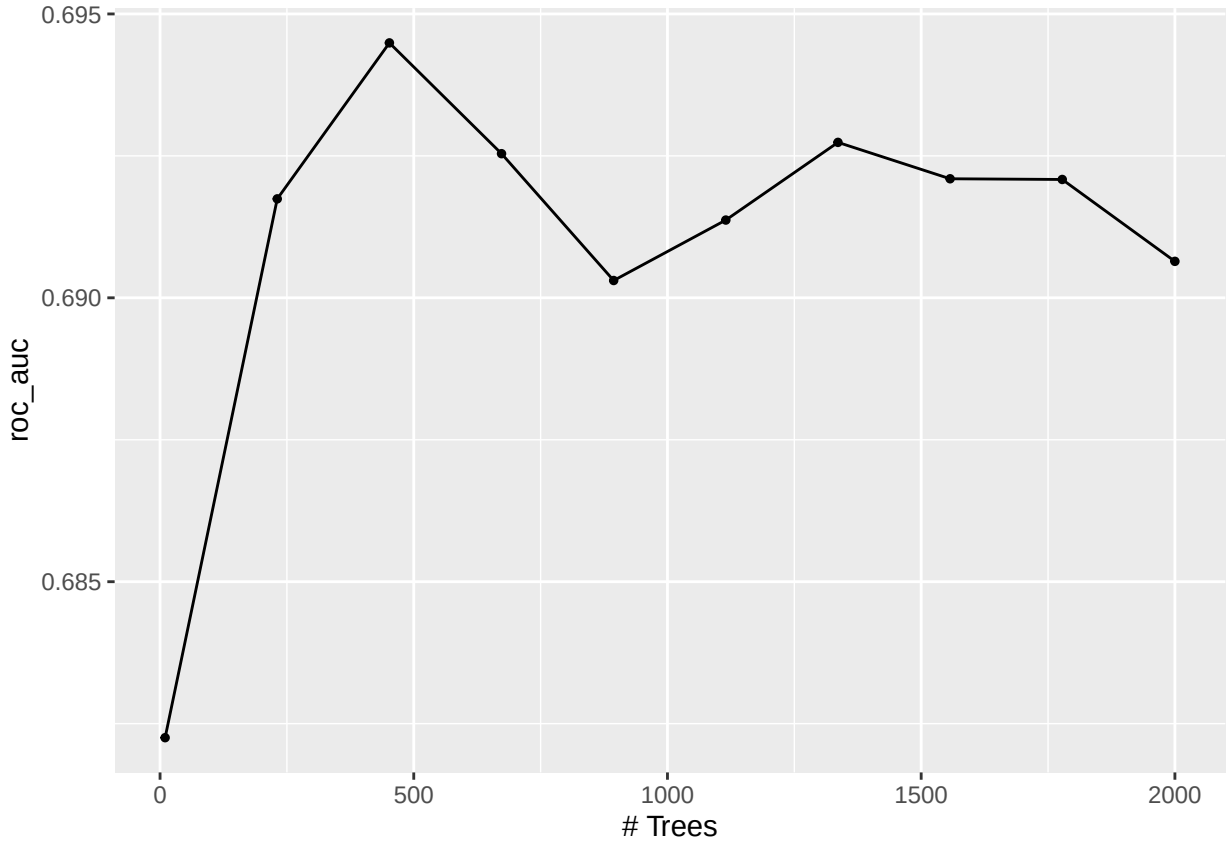
Exercise 9

Finally, set up a boosted tree model and workflow. Use the `xgboost` engine. Tune `trees`. Create a regular grid with 10 levels; let `trees` range from 10 to 2000. Specify `roc_auc` and again print an `autoplot()` of the results. What is the `roc_auc` of your best-performing boosted tree model on the folds? *Hint: Use `collect_metrics()` and `arrange()`.*

The 'roc_auc' of my best-performing boosted tree model on the folds is 0.6944903.

```
boost_spec <- boost_tree(trees = c(10,2000), tree_depth = 4) %>%
  set_engine("xgboost") %>%
  set_mode("classification")

param_grid_boost <- grid_regular(trees(range = c(10, 2000)),  levels = 10)
```

```
boost_workflow <- workflow() %>%
  add_model(boost_spec %>% set_args(trees = tune())) %>%
  add_recipe(pokemon_recipe)

tune_res_boost <- tune_grid(
  boost_workflow,
  resamples = pokemon_fold,
  grid = param_grid_boost,
  metrics = metric_set(roc_auc)
)

autoplot(tune_res_boost)
```




```
collect_metrics(tune_res_boost)
```

```
## # A tibble: 10 x 7
##    trees .metric .estimator  mean     n std_err .config
##    <int> <chr>   <chr>      <dbl> <int>   <dbl> <chr>
##  1    10 roc_auc hand_till  0.682    10  0.0137 Preprocessor1_Model01
##  2   231 roc_auc hand_till  0.692    10  0.0150 Preprocessor1_Model02
##  3   452 roc_auc hand_till  0.694    10  0.0146 Preprocessor1_Model03
##  4   673 roc_auc hand_till  0.693    10  0.0147 Preprocessor1_Model04
##  5   894 roc_auc hand_till  0.690    10  0.0154 Preprocessor1_Model05
##  6  1115 roc_auc hand_till  0.691    10  0.0156 Preprocessor1_Model06
##  7  1336 roc_auc hand_till  0.693    10  0.0157 Preprocessor1_Model07
```

```
##  8  1557 roc_auc hand_till  0.692    10  0.0157 Preprocessor1_Model08
##  9  1778 roc_auc hand_till  0.692    10  0.0156 Preprocessor1_Model09
## 10  2000 roc_auc hand_till  0.691    10  0.0159 Preprocessor1_Model10
```

```
arrange(tune_res_boost)
```

```
## # Tuning results
## # 10-fold cross-validation
## # A tibble: 10 x 4
##    splits           id     .metrics          .notes
##    <list>           <chr>  <list>            <list>
##  1 <split [306/35]> Fold01 <tibble [10 x 5]> <tibble [0 x 3]>
##  2 <split [307/34]> Fold02 <tibble [10 x 5]> <tibble [0 x 3]>
##  3 <split [307/34]> Fold03 <tibble [10 x 5]> <tibble [0 x 3]>
##  4 <split [307/34]> Fold04 <tibble [10 x 5]> <tibble [0 x 3]>
##  5 <split [307/34]> Fold05 <tibble [10 x 5]> <tibble [0 x 3]>
##  6 <split [307/34]> Fold06 <tibble [10 x 5]> <tibble [0 x 3]>
##  7 <split [307/34]> Fold07 <tibble [10 x 5]> <tibble [0 x 3]>
##  8 <split [307/34]> Fold08 <tibble [10 x 5]> <tibble [0 x 3]>
##  9 <split [307/34]> Fold09 <tibble [10 x 5]> <tibble [0 x 3]>
## 10 <split [307/34]> Fold10 <tibble [10 x 5]> <tibble [0 x 3]>
```

```
best_complexity3 <- select_best(tune_res_boost)
best_complexity3
```

```
## # A tibble: 1 x 2
##   trees .config
##   <int> <chr>
## 1   452 Preprocessor1_Model03
```

Exercise 10

Display a table of the three ROC AUC values for your best-performing pruned tree, random forest, and boosted tree models. Which performed best on the folds? Select the best of the three and use `select_best()`, `finalize_workflow()`, and `fit()` to fit it to the *testing* set.

Print the AUC value of your best-performing model on the testing set. Print the ROC curves. Finally, create and visualize a confusion matrix heat map.

Which classes was your model most accurate at predicting? Which was it worst at? Note that classifiers that give corners closer to the top left means a good prediction. Therefore, classes "Bug", "Fire", "Grass", "Normal" were the most accurate at predicting and "Psychic" and "Water" were the worst at predicting.

```
set.seed(1)
df <- data.frame(best_performing = c(0.6683159, 0.7219057, 0.6944903),
                 models <- c("pruned tree model", "random forest model", "boosted tree model"))

head(df)
```

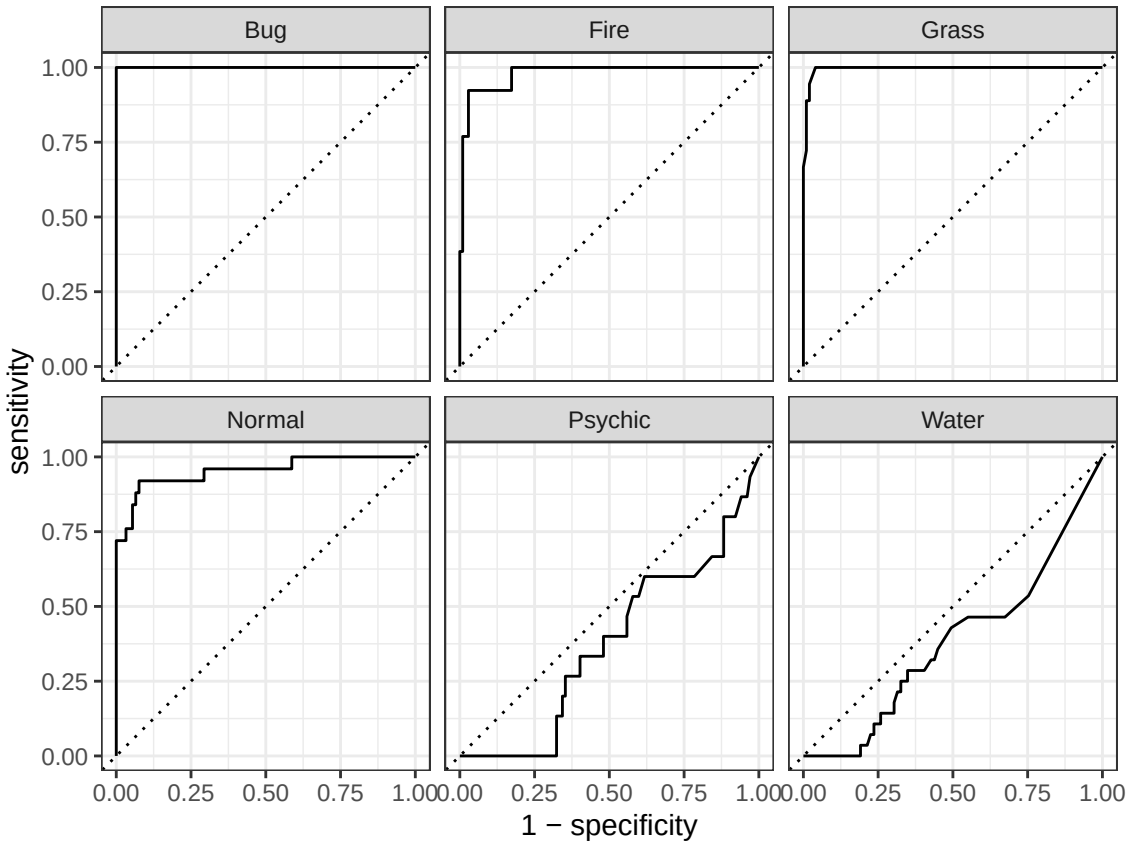
```
##   best_performing
## 1       0.6683159
## 2       0.7219057
```

```
## 3       0.6944903
##   models....c..pruned.tree.model....random.forest.model....boosted.tree.model..
## 1                                                                 pruned tree model
## 2                                                               random forest model
## 3                                                                boosted tree model
```

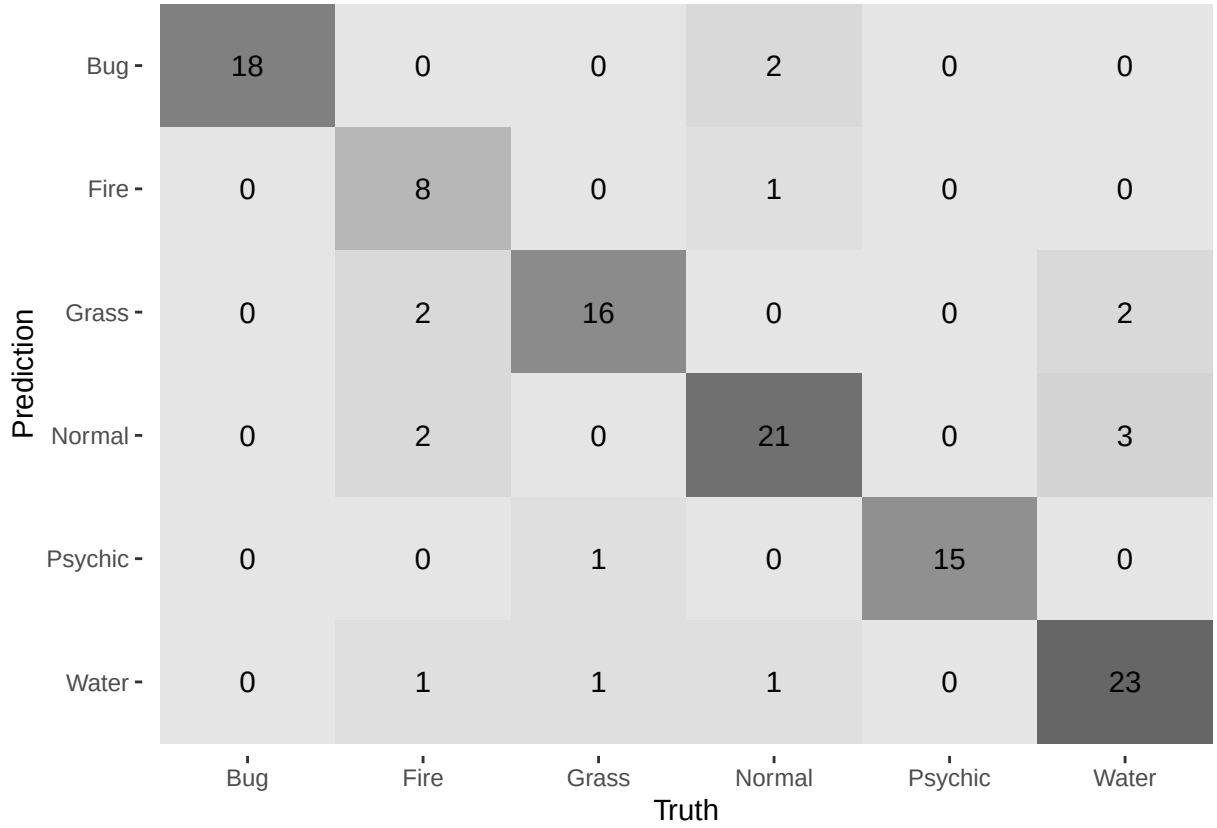
```
#fit it to the testing set
best_complexity <- select_best(tune_res)

class_tree_final <- finalize_workflow(forest_workflow, best_complexity2)

class_tree_final_fit <- fit(class_tree_final, data = pokemon_test)

#AUC value
pred_result <- augment(class_tree_final_fit, new_data = pokemon_test)
auc <- roc_auc(data = pred_result, truth = type_1, estimate = c(.pred_Bug, .pred_Fire, .pred_Grass, .pr
auc
```

```
## # A tibble: 1 x 3
##   .metric .estimator     .estimate
##   <chr>   <chr>              <dbl>
## 1 roc_auc macro_weighted     0.758
```

```
#roc curve
augment(class_tree_final_fit, new_data = pokemon_test) %>%
  roc_curve(type_1, estimate = .pred_Bug, .pred_Fire, .pred_Grass, .pred_Normal, .pred_Water, .pred_Psy
  autoplot()
```




```
#confusion matrix heat map
augment(class_tree_final_fit, new_data = pokemon_test) %>%
  conf_mat(truth = type_1, estimate = .pred_class) %>%
  autoplot(type = "heatmap")
```




```
# best_performing <- c(best_complexity, best_complexity2, best_complexity3)
# models <- c("pruned tree model", "random forest model", "boosted tree model")
# results <- tibble(best_model = best_performing, models = models)



#AUC value
pred_result <- augment(class_tree_final_fit, new_data = pokemon_test)
auc <- roc_auc(data = pred_result, truth = type_1, estimate = c(.pred_Bug, .pred_Fire, .pred_Grass, .pr
auc
```

```
## # A tibble: 1 x 3
##   .metric .estimator     .estimate
##   <chr>   <chr>              <dbl>
## 1 roc_auc macro_weighted     0.758
```