

## 1) Descrição do problema:

O mundo Pokémon é vasto, possuindo centenas de espécies, cada qual com sua peculiaridade. Um dos papéis que o jogador tem nos jogos é de catalogar todas as espécies existentes, tarefa nada fácil. Ao completar os registros, o jogador é recompensado com alguns itens, todavia nenhuma análise mais aprofundada sobre eles é feita. Este trabalho irá analisar um conjunto de dados contendo informações de 721 Pokémon para buscar compreender mais a fundo as características principais deles.

Um Pokémon possui 6 atributos: vida (HP), ataque, (Attack), defesa (Defense), ataque especial (SP. Attack), defesa especial (SP. Defense), velocidade (Speed). Comumente um sétimo atributo é levado em conta para análises individuais, o total (Total), sendo ele a soma dos outros sete.

Nos últimos anos, o jogo tomou uma característica competitiva e, devido à diversidade de espécies, se tornou um desafio selecionar quais são as melhores para se utilizar. Um Pokémon será geralmente considerado bom caso ele satisfaça alguma das seguintes condições:

- Alta velocidade (geralmente mais de 100)
- Alta quantidade de atributos totais
- Possuir algum atributo acima da média
- Combinação de tipagem única

Em casos raros, possuir um peso alto é considerado uma vantagem, visto que certos golpes o levam em consideração para o cálculo de dano.

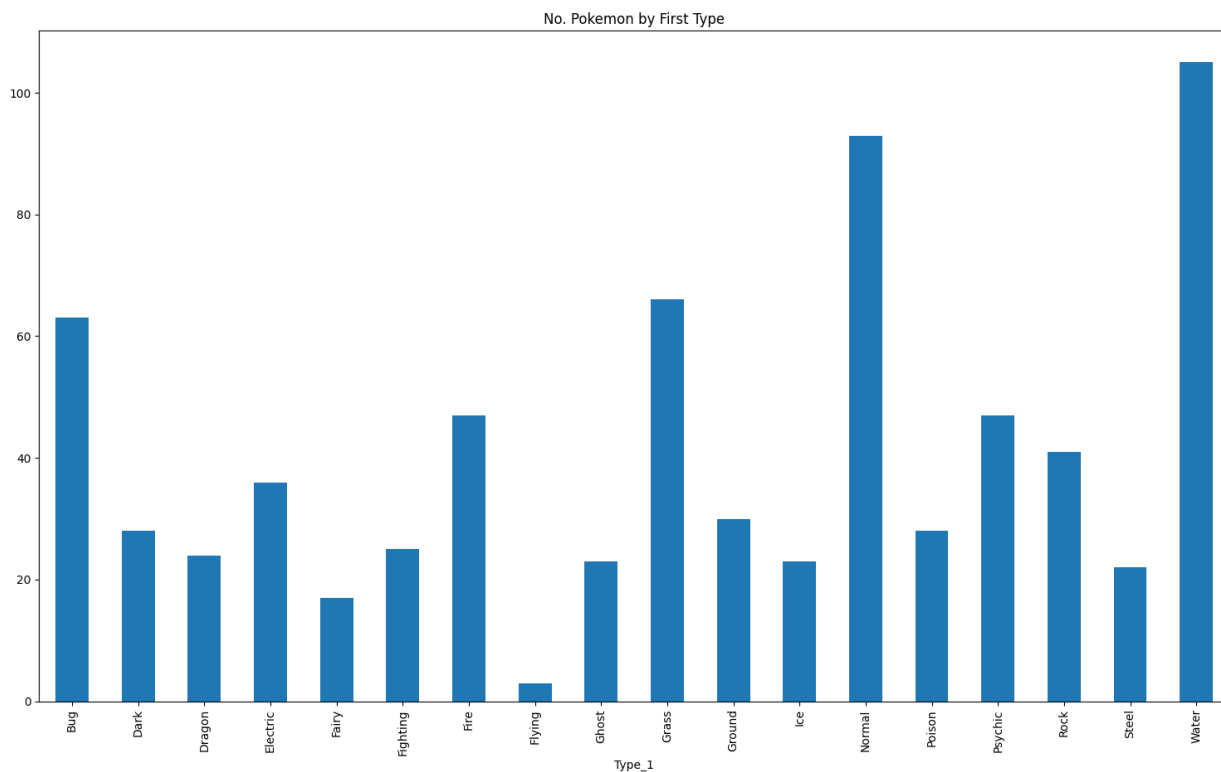
Busca-se encontrar alguma correlação entre a tipagem de um Pokémon e seus atributos, bem como são feitas as distribuições deles e quais combinações de tipo são as mais comuns. Embora não seja fator decisivo, é interessante saber se o peso pode influenciar alguns atributos, como a velocidade e o total

Uma geração de Pokémon é representada pela quantidade de Pokémon que são acrescentados a cada lançamento de jogo. Um dos objetivos desta análise também é descobrir se houve algum aumento significativo nos atributos ao longo das gerações, possibilitando compreender se os novos Pokémon estão involuntariamente trazendo os mais antigos para a obsolescência.

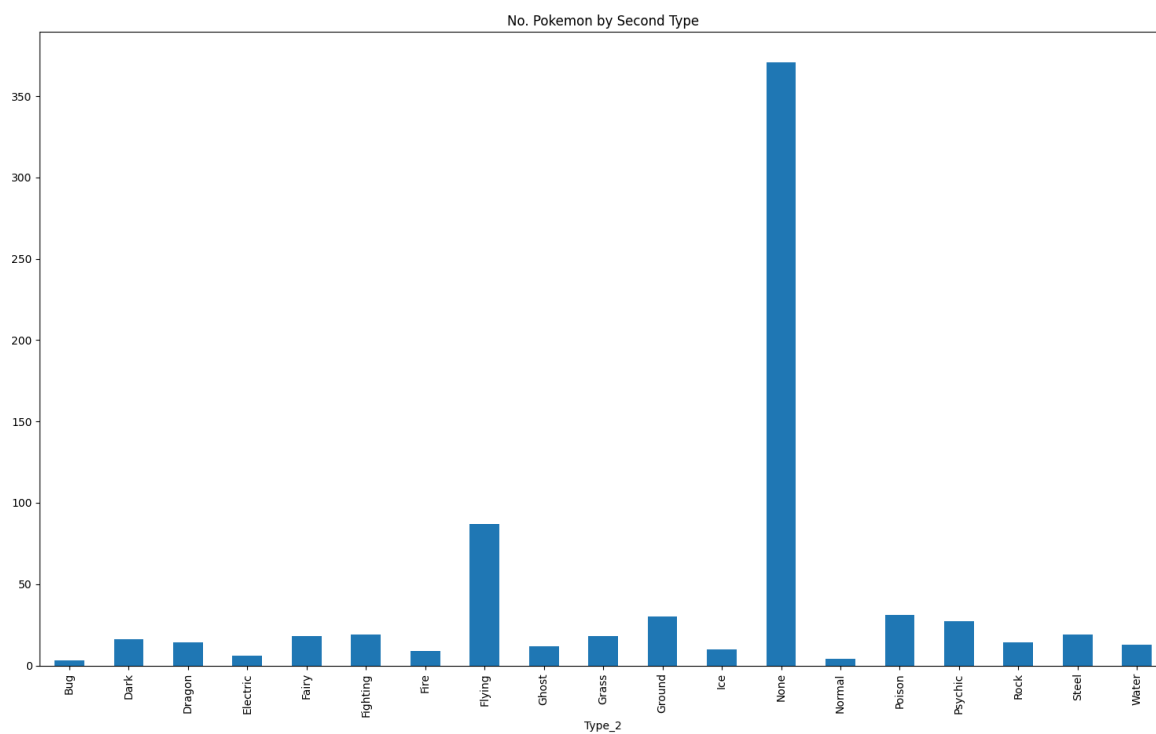
## 2) Questões para uma análise individual de cada variável:

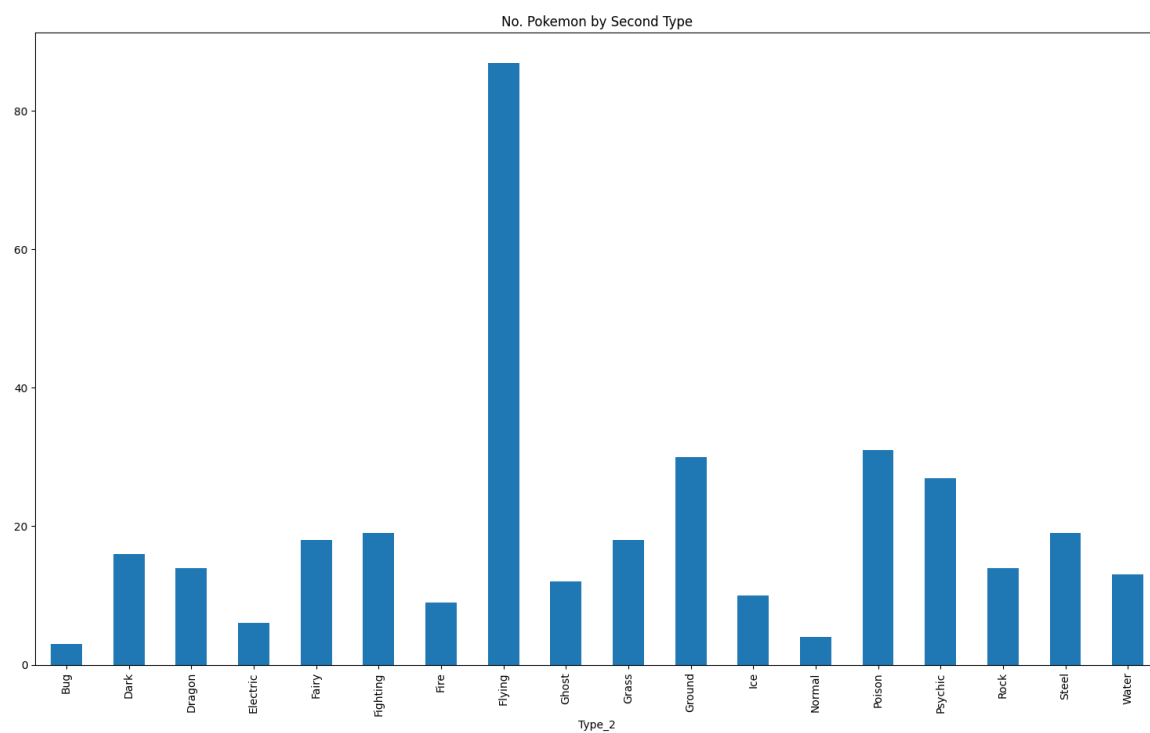
A variável *Name* (nome) não foi utilizada pois ela é unicamente um identificador.

Variável *Type\_1* (tipo 1): quantos Pokémon há para cada tipo primário?

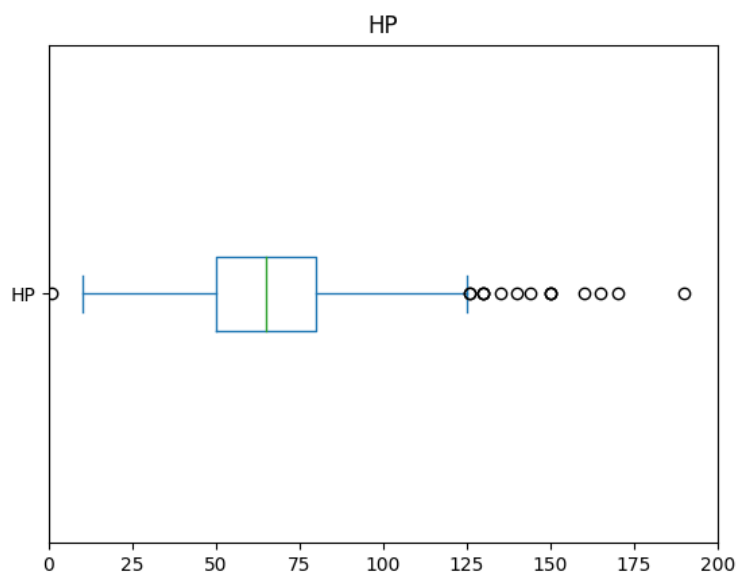


Variável *Type\_2* (tipo 2): quantos Pokémon há para cada tipo secundário? Dois gráficos foram feitos, um deles incluindo os Pokémon que não possuem segundo tipo.

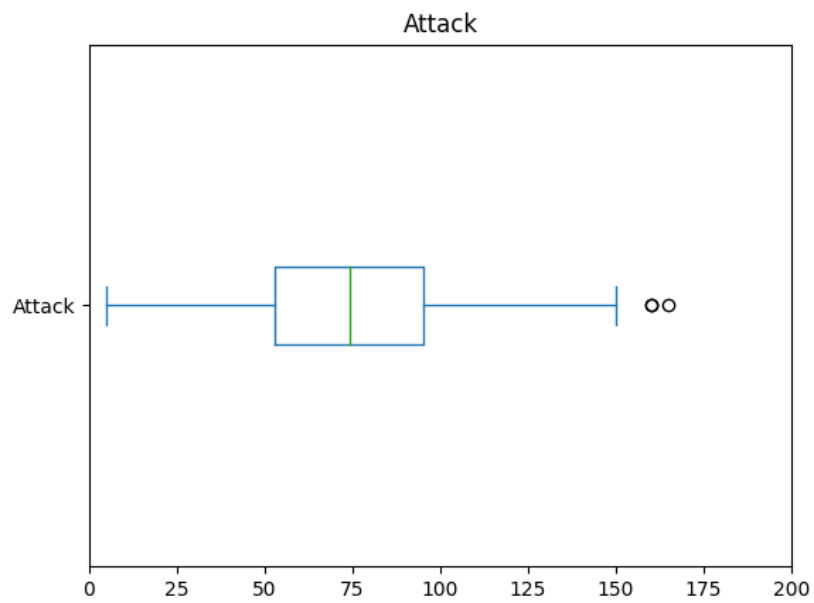




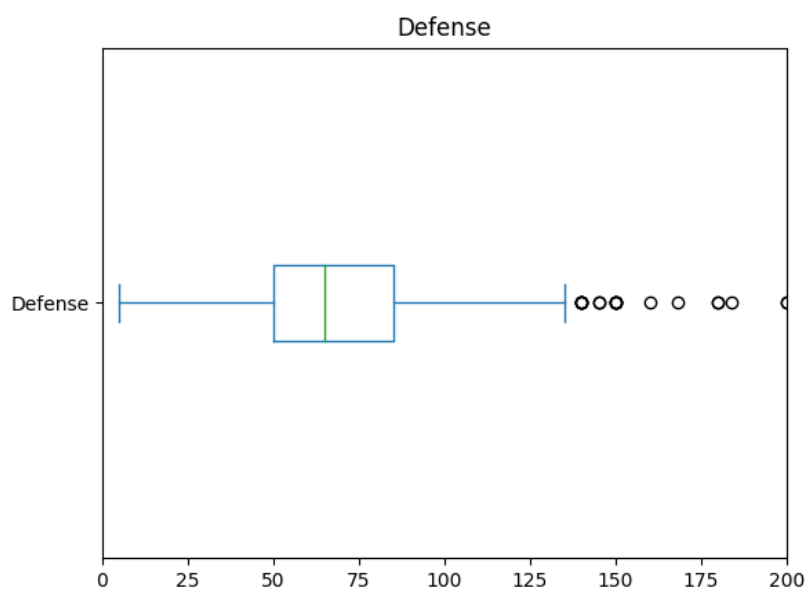
Variável *HP* (vida): qual a distribuição do atributo?



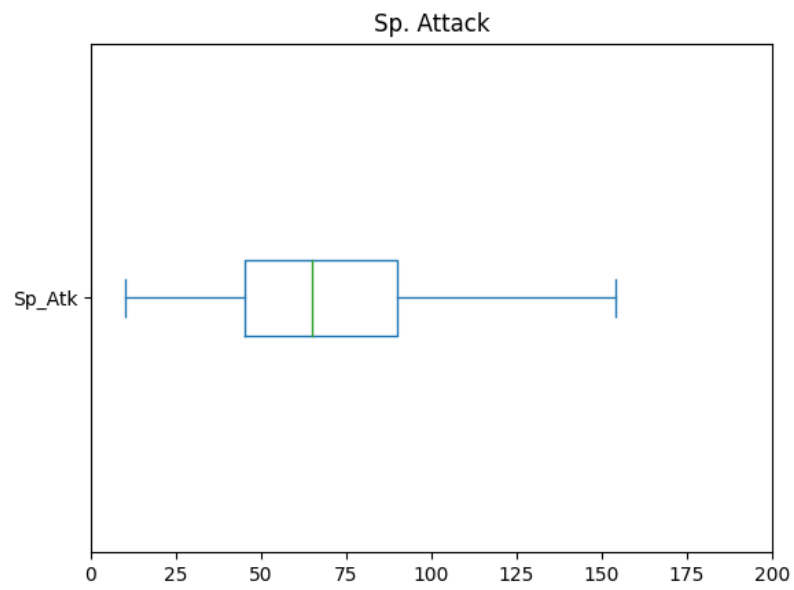
Variável *Attack* (ataque): qual a distribuição do atributo?



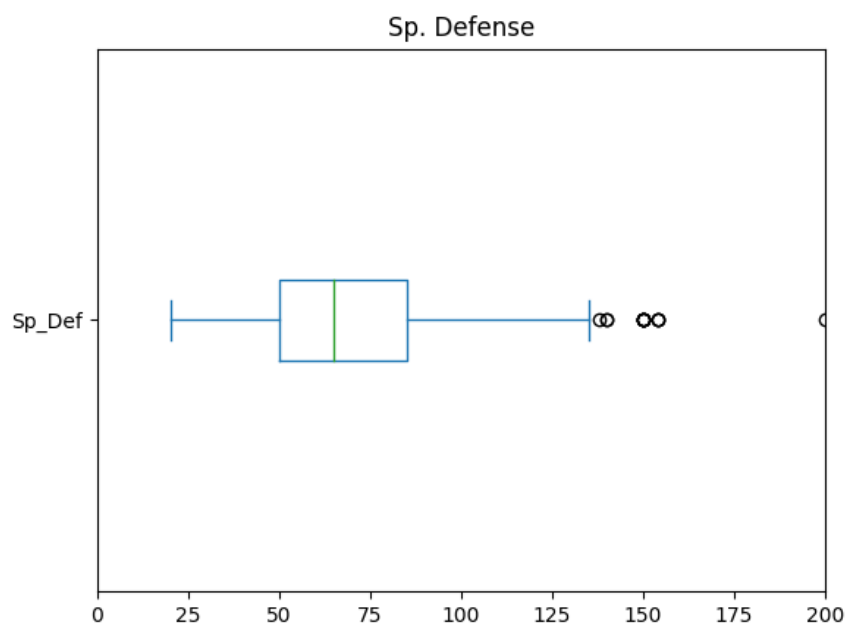
Variável *Defense* (defesa): qual a distribuição do atributo?



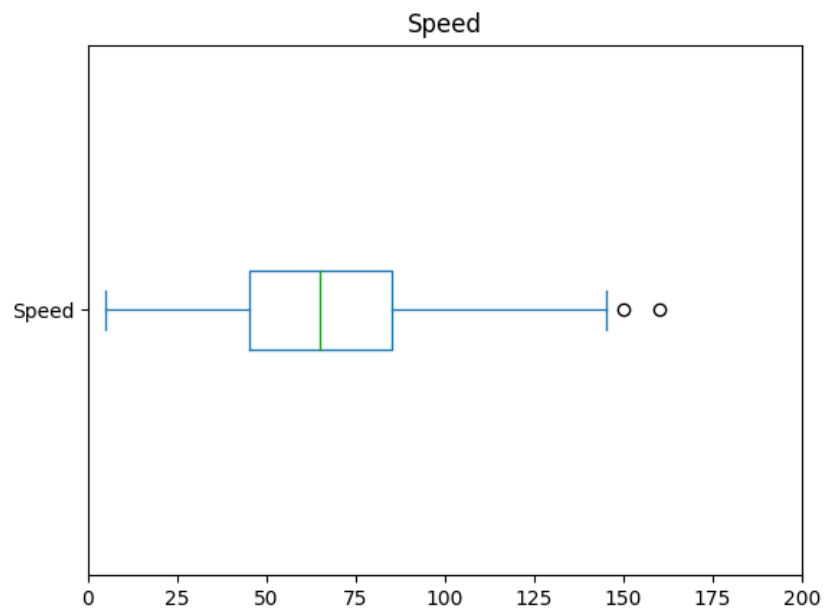
Variável *Sp\_Atk* (ataque especial): qual a distribuição do atributo?



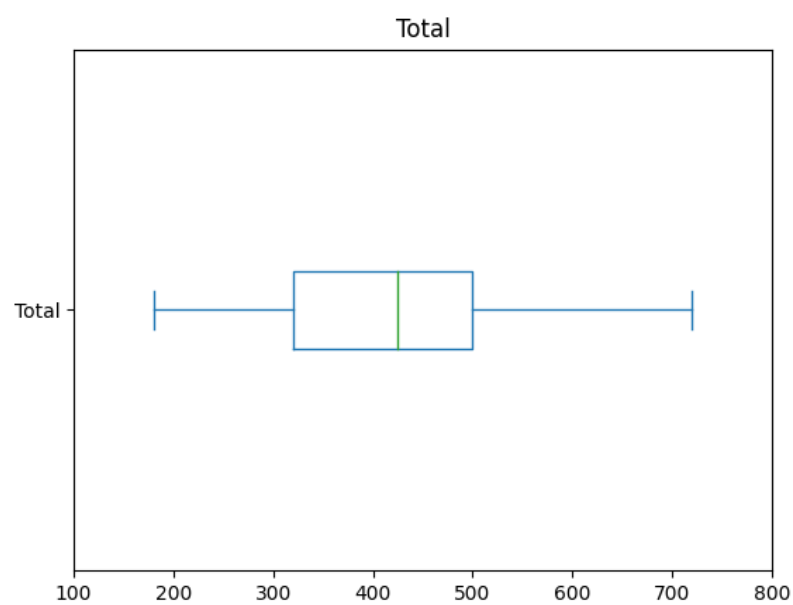
Variável *Sp\_Defense* (defesa especial): qual a distribuição do atributo?



Variável *Speed* (velocidade): qual a distribuição do atributo?

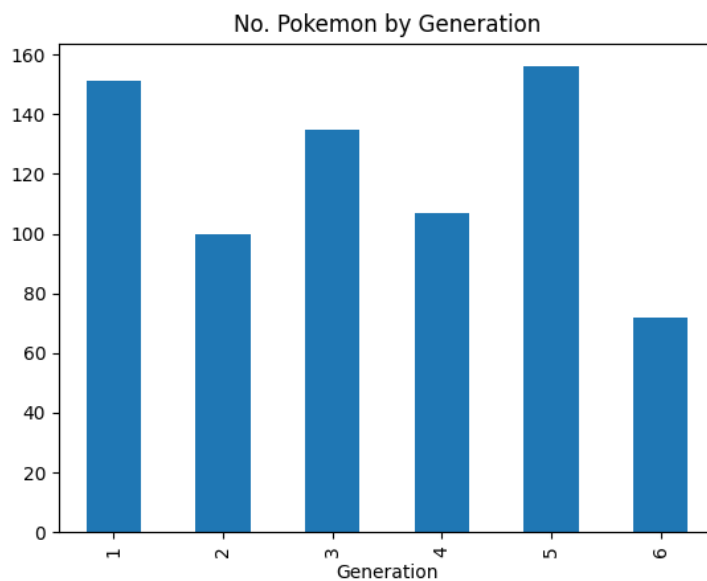


Variável *Total* (total): qual a distribuição do atributo?

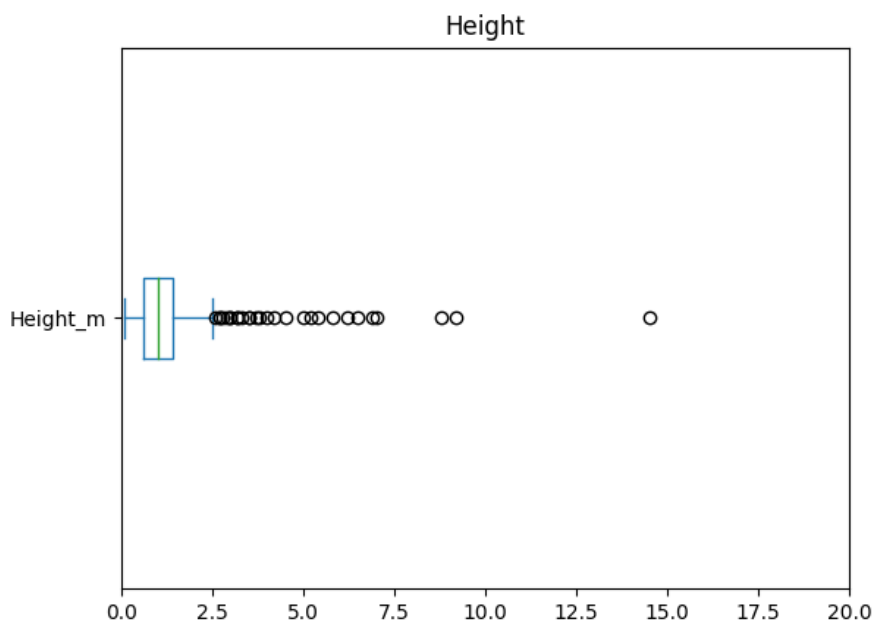


]

Variável *Generation* (geração): quantos Pokémon há por geração?

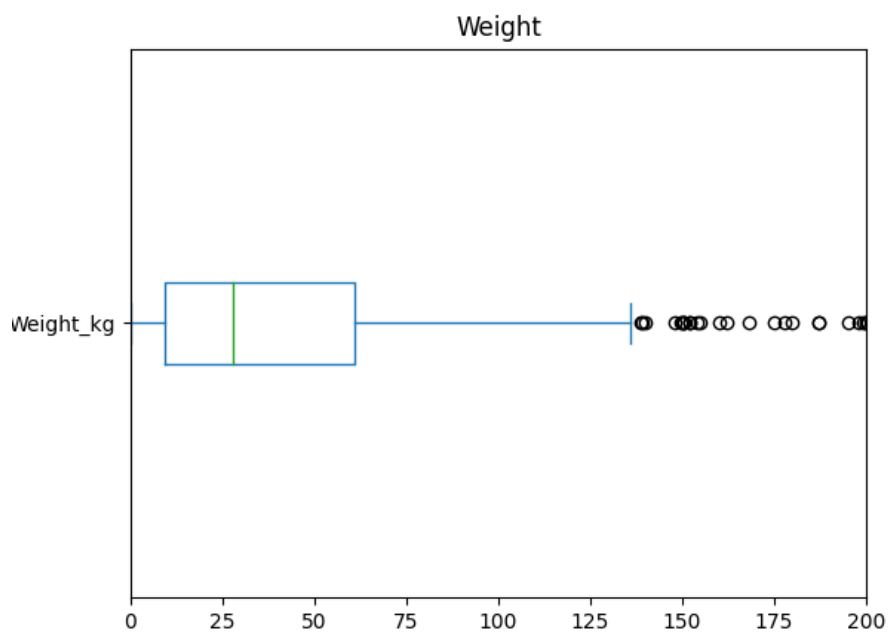


Variável *Height\_m* (altura): embora comumente vejamos Pokémon grandes na mídia, como dragões, eles são realmente altos em sua maioria?



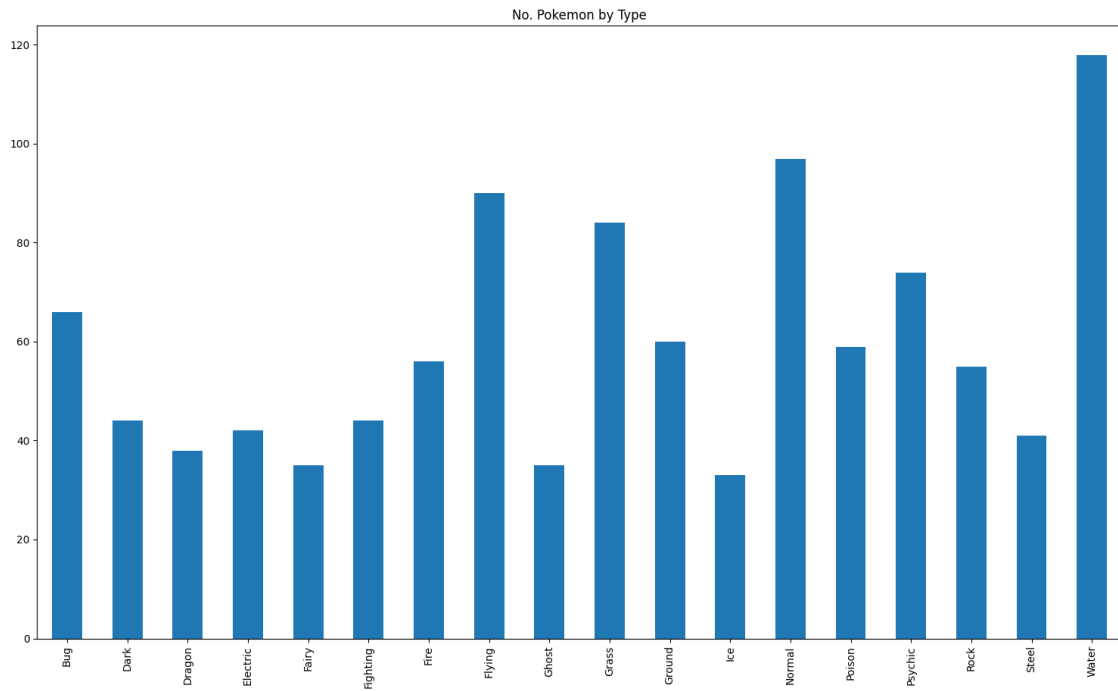


Variável *Weight\_kg* (peso): os Pokémon são criaturas naturalmente pesadas/leves ou seu peso pode variar?

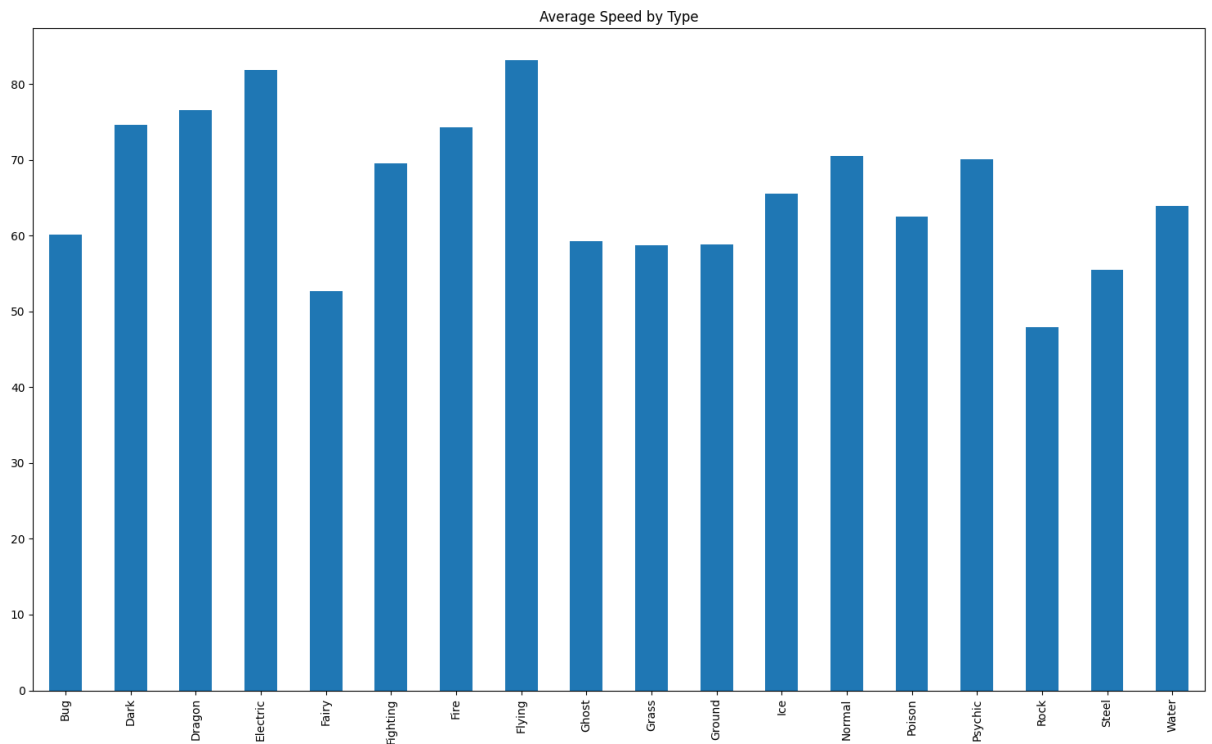


### 3) Questões para uma análise individual das relações entre as variáveis:

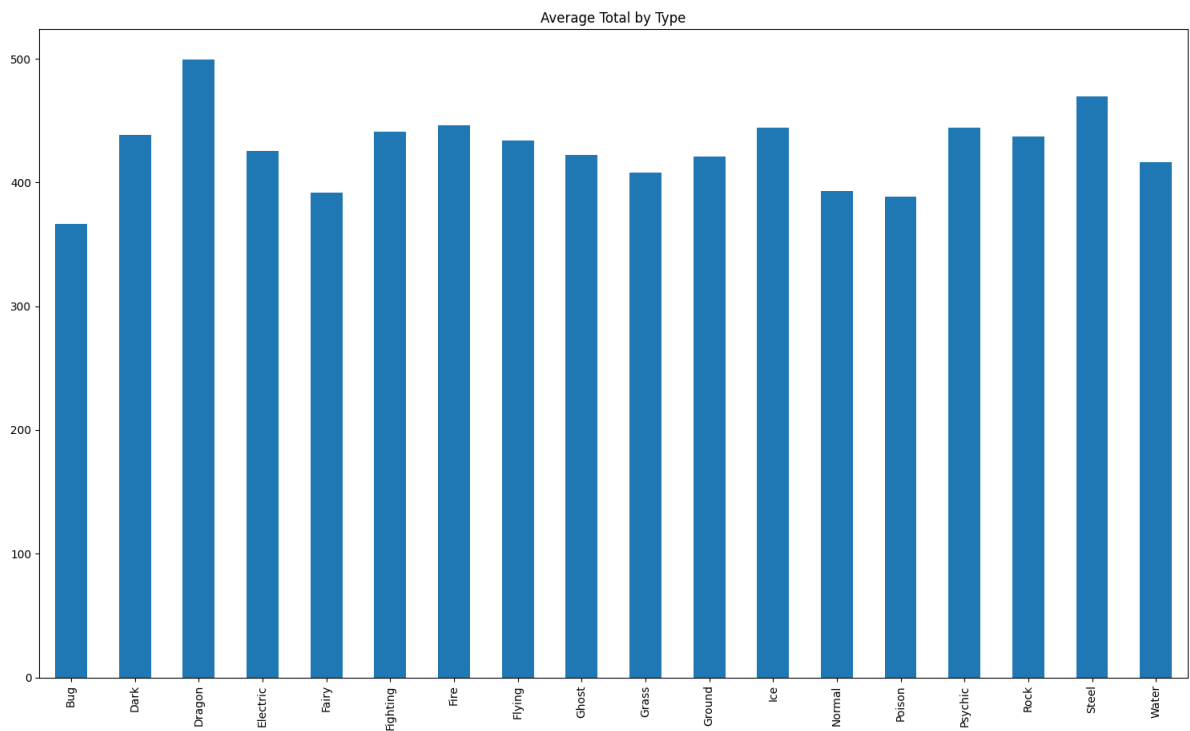
Se somarmos as ocorrências de tipagens (seja primária ou secundária), teremos uma distribuição mais uniforme entre elas?



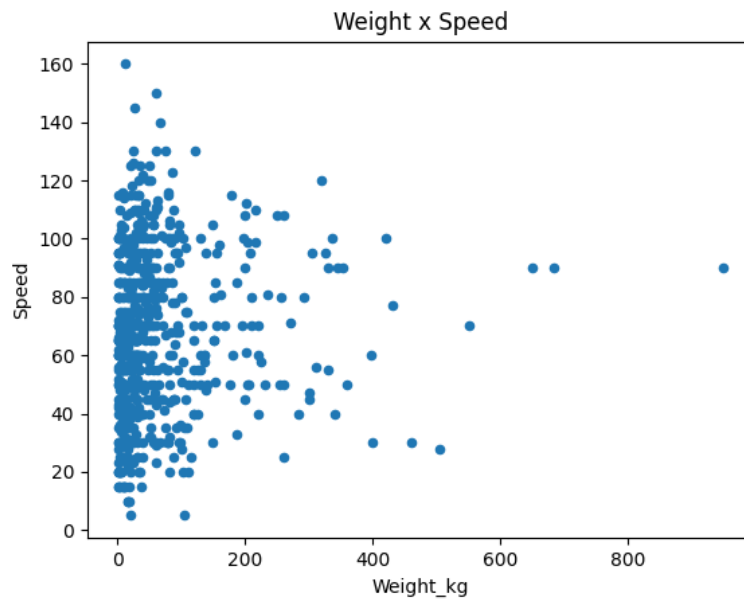
Existem tipos mais propensos a serem mais rápidos ou lentos?



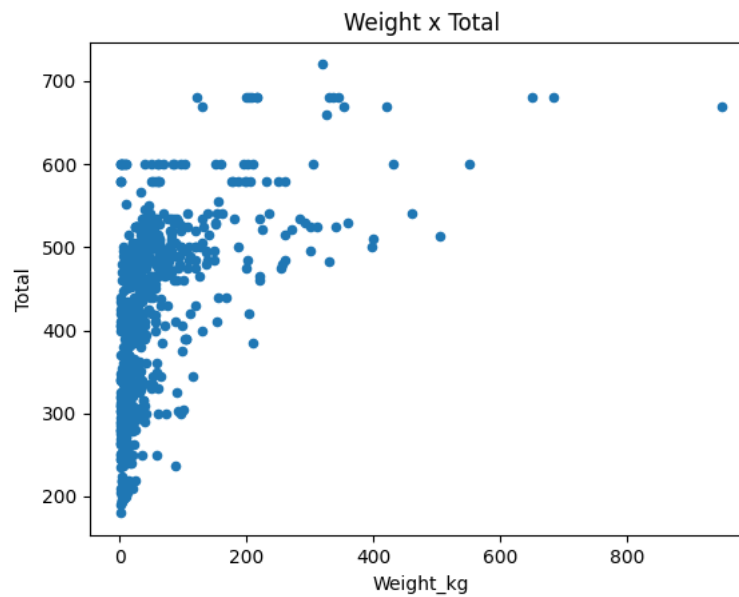
Existem tipos mais fortes (maior valor total) que outros?



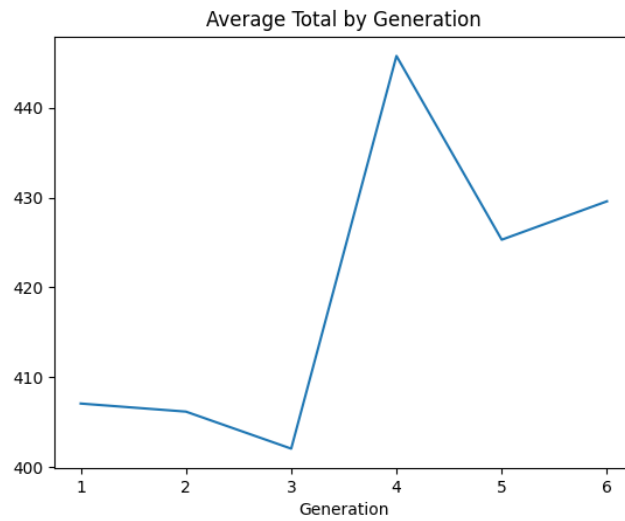
Seriam os Pokémon mais pesados mais lentos?



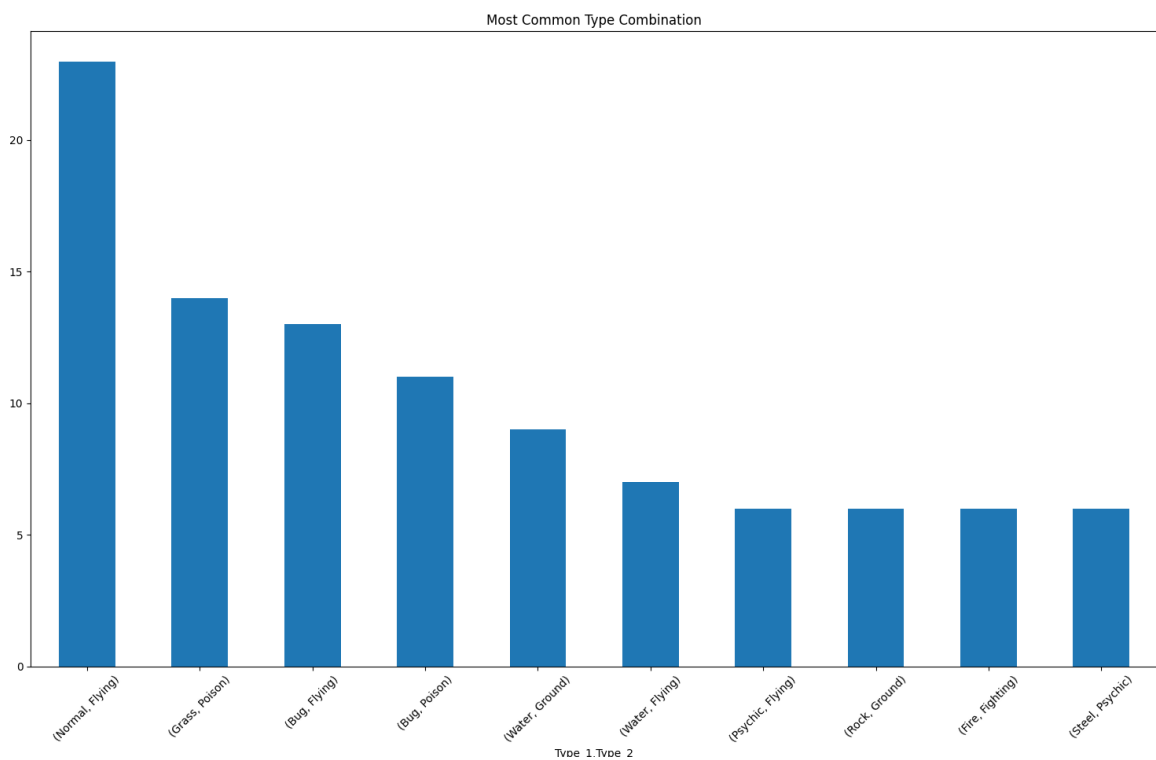
Seriam os Pokémon mais pesados mais fortes em atributos totais?



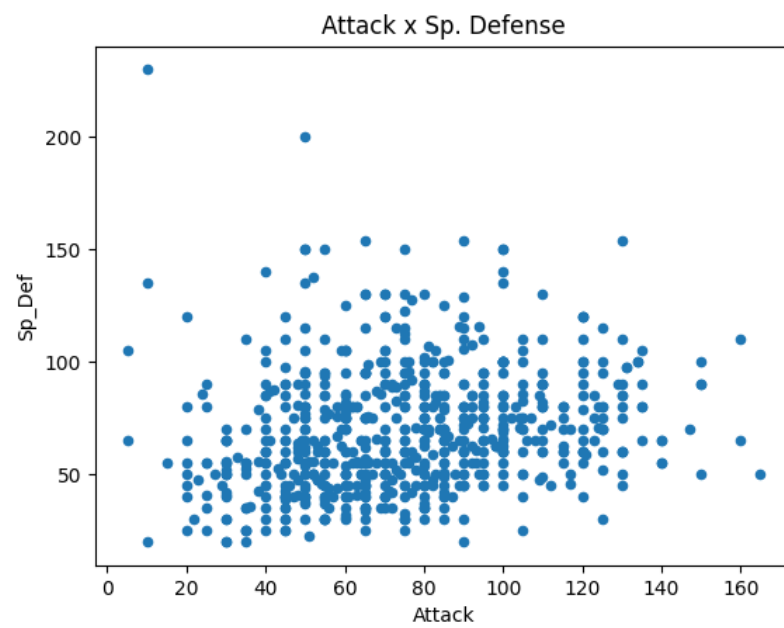
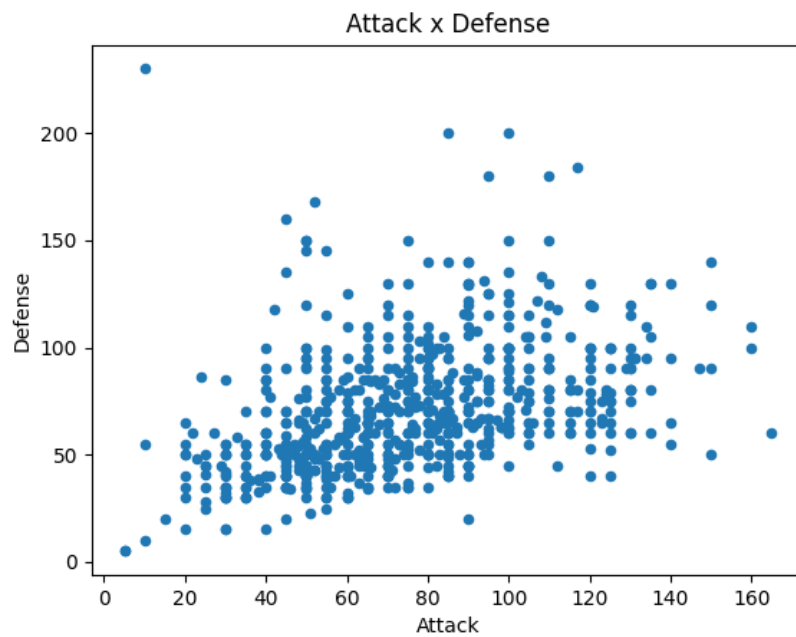
As gerações mais recentes introduziram espécies mais fortes (atributos totais)?



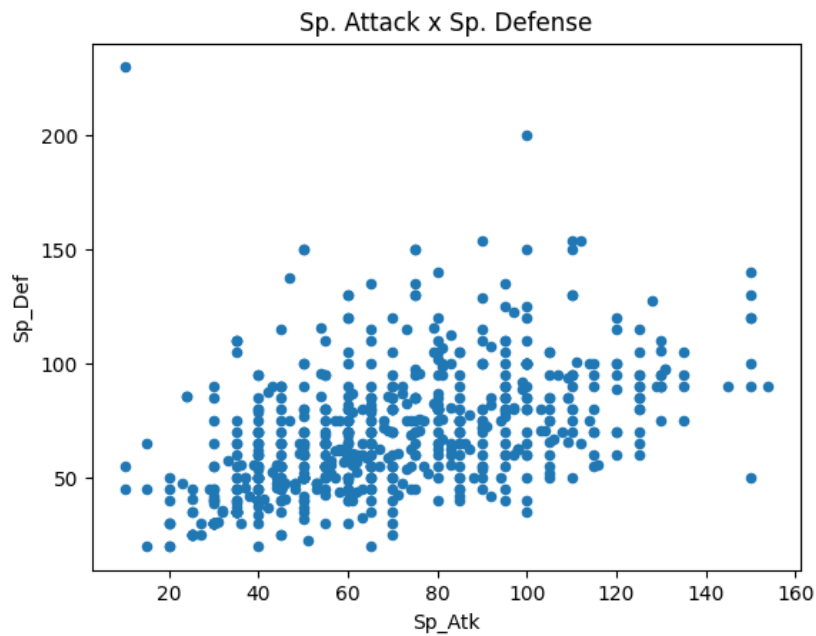
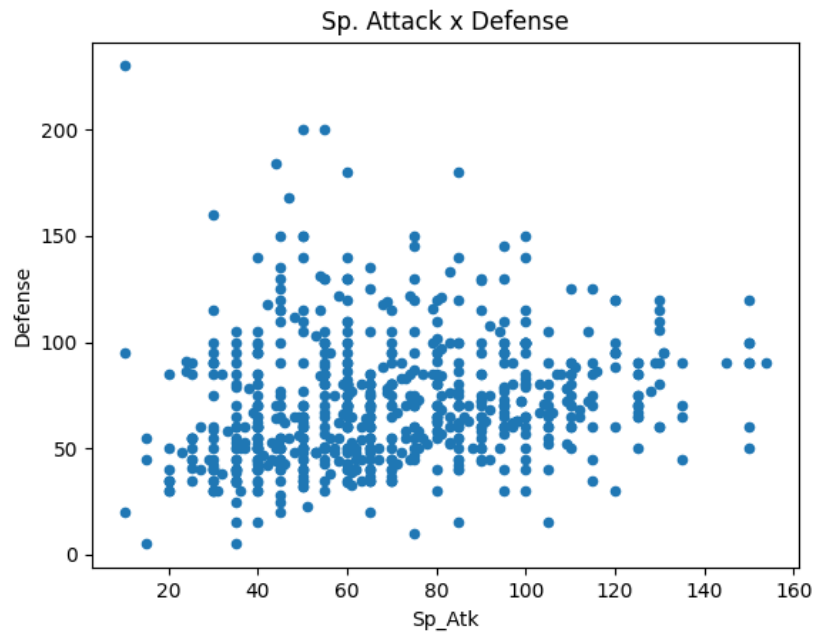
Quais são as combinações de tipos mais comuns?



Seriam Pokémon com maior ataque menos resistentes (menor defesa/defesa especial)?



Seriam Pokémon com maior ataque especial menos resistentes (menor defesa/defesa especial)?



#### 4) Insights levantados da análise exploratória preliminar realizada na base:

Percebe-se que os atributos ofensivos (ataque/ataque especial) contemplam um alcance maior, na faixa de ~ 5-150, que os defensivos (defesa/defesa especial), na faixa de ~ 5-130.

É mais comum existirem *outliers* para atributos defensivos que os ofensivos.

Não parece haver uma correlação muito significativa entre maiores ataques com menores defesas, as faixas possuem um tamanho bem parecido ao longo de todo o gráfico.

A partir da quarta geração a quantidade de atributos totais passou por uma leve inflada, todavia não maior que ~10%. Isso pode ser um indicativo do fenômeno *power creep*, o qual se caracteriza pelo desbalanceamento gradual de jogos devido ao lançamento de conteúdos novos que colocam os mais antigos como mais fracos.

Os tipos não são igualmente distribuídos. Mesmo com alguns aparecendo mais/menos frequentemente como primário ou secundário.

Os Pokémon mais pesados aparentam ter uma velocidade que esteja na média dos mais leves, mas não é possível afirmar devido às pequenas amostras.

Os Pokémon mais pesados aparentam ter uma quantidade total de atributos significativamente maior que a média, mas não é possível afirmar devido às pequenas amostras.

Embora existam Pokémon muito pesados e altos, quase todos não passam dos 2,5 metros e ~130 quilos.

Existem tipos que são mais velozes, bem como mais lentos, em média. Há uma grande discrepância entre tipos rápidos e lentos.

Diferentemente do gráfico de velocidade, os atributos totais são melhores distribuídos entre os tipos, ainda que existam algumas diferenças como entre Bug e Dragon.

Com exceção das combinações de tipagem Normal/Flying, Grass/Poison, Bug/Flying, Bug/Poison e Water/Ground, as outras aparentam ser melhores distribuídas (importante notar que o gráfico não contempla todas, visto que há milhares possíveis).