Data Scientist - Take Home Challenge

Kyeongho Min | [longmini.seo@gmail.com](mailto:longmini.seo@gmail.com) | (+1) 408-609-0254

## Question 1

Suppose this data is a SQL table called ‘PokemonStats’. In an SQL dialect you are most comfortable with, find the top 3 Pokemon in terms of total stats of each type (primary type, Type\_1). Your answer should include: 1) the SQL dialect you are using; 2) The SQL query used to answer this question; 3) The returned result.

1. The dialect I used is MySQL
2. SQL query:

|  |
| --- |
| SELECT p.Type\_1, p.Name, p.Total  FROM PokemonStats as p  WHERE(SELECT count(\*) FROM PokemonStats as k  WHERE k.Type\_1 = p.Type\_1 and k.Total >= p.Total) <= 3  ORDER BY p.Type\_1 ASC, p.Total DESC |

1. The returned result

|  |
| --- |
| Type\_1 Name Total  0 Bug Genesect 600  1 Bug Volcarona 550  2 Bug Yanmega 515  3 Dark Yveltal 680  4 Dark Darkrai 600  5 Dark Hydreigon 600  6 Dragon Rayquaza 680  7 Dragon Reshiram 680  8 Dragon Zekrom 680  9 Electric Zapdos 580  10 Electric Raikou 580  11 Electric Thundurus 580  12 Fairy Xerneas 680  13 Fairy Florges 552  14 Fairy Togekiss 545  15 Fighting Lucario 525  16 Fighting Mienshao 510  17 Fire Ho-Oh 680  18 Fire Heatran 600  19 Fire Volcanion 600  20 Flying Tornadus 580  21 Flying Noivern 535  22 Flying Noibat 245  23 Ghost Giratina 680  24 Ghost Dusknoir 525  25 Ghost Chandelure 520  26 Grass Shaymin 600  27 Grass Virizion 580  28 Grass Tangrowth 535  29 Ground Groudon 670  30 Ground Landorus 600  31 Ground Rhyperior 535  32 Ice Articuno 580  33 Ice Regice 580  34 Ice Vanilluxe 535  35 Normal Arceus 720  36 Normal Slaking 670  37 Normal Regigigas 670  38 Poison Crobat 535  39 Poison Nidoqueen 505  40 Poison Nidoking 505  41 Psychic Mewtwo 680  42 Psychic Lugia 680  43 Rock Tyranitar 600  44 Rock Diancie 600  45 Steel Dialga 680  46 Steel Metagross 600  47 Steel Jirachi 600  48 Water Palkia 680  49 Water Kyogre 670  50 Water Manaphy 600 |

## ***For the next 3 questions below, I would like to answer the questions via Jupyter notebook with additional .ipynb file. OR, I posted my notebook file on github, please take a look at it if you want. Thank you***

Link : <https://github.com/longmini0/Niantic_challenge/blob/main/pokemon_challenge_Q2-Q4.ipynb>

## Question 2

Imagine a new Pokemon game where you are only allowed to collect ONE type of Pokemon. Similar to other Pokemon games, your goal is to have the strongest battlers and defenders for battles and raids. Which type will you pick? Why?

## Question 3

If you want to predict whether the Pokemon is a legendary Pokemon (a.k.a. predict the field isLegendary using other fields), which models would you use? List your top 3 models with pros and cons for each one.

## Question 4

Pick one model and implement it in a language you are most comfortable with (preferably Python or R). Please do not use the ‘Catch\_Rate’ field (if you are Pokemon fan you know why :). What is your in-sample classification accuracy and what fields did you end up using?

Your answer should include: 1) The code of implementing the model (incl. feature processing, model fitting and cross validating); 2) The formula/description of your final model along with the accuracy number. 3) In addition to the code and the model specification, if you choose to submit a presentation/ dashboard as part of your writeup, you can present your results in any way you like.