

# Penerapan Data Science Pada Dataset Pokemon

Jonathan Gracio Susanto<sup>#1</sup>, Setia Budi<sup>\*2</sup>

<sup>#</sup>Program Studi Sistem Informasi, Universitas Kristen Maranatha  
Jl. Surya Sumantri No.65 Bandung 40164, Indonesia

<sup>1</sup>1873022@maranatha.ac.id

<sup>2</sup>setia.budi@it.maranatha.edu

**Abstract** — Data science is a science that studies data. Data that is usually explored in data science is quantitative data. In short, quantitative data is data that can be calculated because the data is in the form of numbers. Required knowledge of data science is mathematics, statistics and programming such as python. In this case, the dataset used is the pokemon dataset. This pokemon dataset has some basic information such as name, type, generation, HP, attack, defense, speed and legendary. This dataset is obtained from the kaggle website, this kaggle website is a website for the data science community, on this website everyone can share the dataset and even some of the exploration code they have done so that other people can see it on the kaggle website. There are several potential explorations that can be done from this pokemon dataset, including analyzing the data on the pokemon dataset, making data visualizations and making the exploration results into new information that is more useful than before. This exploration is assisted by several tools, namely python pandas and jupyter notebook.

**Keywords**— data science, kaggle, pokemon, python, data visualization

## I. PENDAHULUAN

Pokemon adalah salah satu seri *game* yang diciptakan di Jepang oleh Satoshi Tajiri. Pokemon ini merupakan *game* yang bergenre *game* petualangan. *Game* Pokemon ini awalnya berada pada *game console* yang disebut *Game Boys*, lalu Nintendo dan sampai pada *game mobile*. Seiring berjalannya waktu semakin banyak *game* pokemon yang dirilis. Pokemon sendiri itu memiliki beberapa informasi seperti nama, tipe dan kemampuan yang berbeda beda. Kemampuan ini meliputi *attack*, *defense*, *special ability* dan *speed* dari pokemon.

Untuk mengetahui lebih dalam informasi pokemon, maka terdapat sebuah dataset pokemon. Dataset ini memiliki beberapa informasi seperti nama pokemon, *type 1*, *type 2*, *HP*, *attack*, *defense*, *special attack*, *special defense*, *speed*, *generation* dan *legendary*. *Dataset* yang ada ini adalah data dari *game* pokemon tidak termasuk data yang ada pada Pokemon Cards atau Pokemon Go.

Terdapat beberapa potensi eksplorasi yang dapat dilakukan pada *dataset* tersebut yaitu menganalisis tipe pada pokemon, menganalisis pokemon dari setiap generasi, membandingkan pokemon legendary dan non legendary, menganalisis pokemon yang berevolusi, menganalisis top 10 pokemon berdasarkan beberapa *variable* dan merekomendasikan beberapa pokemon. Dari beberapa eksplorasi tersebut diharapkan bisa menghasilkan informasi yang lebih berguna dari sebelumnya.

## II. KAJIAN TEORI

### A. Data

Mengirimkan data saat ini begitu mudah, data dapat dikirimkan melalui internet. Data itu sendiri berisikan informasi yang diperoleh melalui pengamatan atau bisa juga disebut dengan kumpulan beberapa fakta. Terdapat 2 jenis data yang sering digunakan dalam penelitian yaitu data kualitatif dan data kuantitatif. Secara singkat data kuantitatif itu adalah data yang dapat diukur atau dihitung karena data tersebut berupa angka, sedangkan data kualitatif data yang disajikan bukan dalam bentuk angka.[1]

### B. Pandas

Pandas ini menyajikan fitur analisis data dan struktur data yang berlisensi *open source*. Pandas ini dibuat untuk mempermudah pekerjaan dengan data yang terstruktur ini. Tipe struktur data yang biasanya dipakai pada Pandas adalah *Dataframe* dan *Series*[2]. Terdapat beberapa data yang bisa dibaca oleh pandas, salah satunya adalah data csv. Untuk membaca data *csv* ini bisa menggunakan fungsi *read csv* yang sudah tersedia pada pandas[3].

### C. Matplotlib

Matplotlib berguna untuk membuat visualisasi data. Dengan matplotlib ini memungkinkan untuk membuat berbagai macam plot seperti bar plot, box plot, line plot dan lainnya. Matplotlib sudah terintegrasi dengan aplikasi yang dinamakan jupyter notebook [4]. Terdapat beberapa manfaat dari matplotlib yaitu [5] [6] :

1. Matplotlib ini mudah digunakan, dengan syntax `import matplotlib as plt` untuk mengimpor *plotting* API yang berguna untuk membuat berbagai macam *plotting*.
2. Matplotlib ini sudah terintegrasi dengan data *analytic packages* yaitu *pandas* dan *numpy*.
3. Bisa mengatur *style* ataupun warna pada plot sehingga data yang ditampilkan bisa lebih representatif.

### D. Numpy

Terdapat library python yang berfungsi untuk melakukan proses komputasi numerik yang disebut Numpy [7]. Numpy ini menyimpan datanya dalam bentuk *array* bisa berupa *array* 1 dimensi atau *array* 2 dimensi. Terdapat beberapa fungsi numpy sebagai berikut [8] :

1. Numpy ini bisa digunakan untuk membuat *array* dengan fungsi *arrange* (`arrange(x)`). X itu mengartikan jumlah *array* yang ingin dibentuk.
2. Numpy juga dapat melakukan operasi numerik seperti penjumlahan, pengurangan, pangkat, perkalian dan lainnya.
3. Numpy ini juga memungkinkan untuk mengecek tipe data yang ada pada suatu *array* dengan suatu fungsi yang disebut *type*.

### E. Jupyter Notebook

*Tool* yang memungkinkan seseorang untuk melakukan analisis dan mengolah data disebut jupyter notebook. Jupyter notebook dapat digunakan untuk menulis kode, menjalankan kode yang dituliskan dan melihat hasil dari kode yang dijalankan dalam satu tampilan yang interaktif. *Output* dari kode yang *diinput* bisa berupa teks sederhana ataupun dalam bentuk grafik. [9]

### F. Exploratory Data Analysis

Exploratory Data Analysis adalah suatu proses eksplorasi terhadap suatu data untuk menemukan hasil atau pola tertentu menggunakan grafik visual atau analisis statistik. EDA juga bisa membantu memahami hubungan antara *variable* yang ada, menemukan informasi yang penting dalam data dan menganalisis suatu pola tertentu. Hal ini akan berguna untuk mendapatkan hasil analisa yang mendalam dan mendapatkan informasi baru yang lebih berguna. [10]

### G. Pokemon

Pocket monster atau yang lebih sering dikenal dengan pokemon diciptakan di Jepang. Pokemon merupakan serangkaian permainan yang dikembangkan oleh Game Freak dan diterbitkan oleh Nintendo. Game pokemon pertama dirilis pada tahun 1996 di Jepang untuk Game Boy, lalu game pokemon ini terus di produksi untuk setiap generasi game console Nintendo, oleh sebab itu pada *dataset* juga terdapat data tentang generasi pokemon. Pada pokemon terdapat beberapa data seperti nama, tipe, *attack*, *defense* yang berguna bagi setiap pemainnya. [11]

## III. METODOLOGI

### A. Dataset Pokemon

*Dataset* yang digunakan dalam eksplorasi ini adalah dataset pokemon. Dataset ini memuat beberapa informasi pokemon yaitu nama, *type 1*, *type 2*, total, HP, *attack*, *defense*, *special attack*, *special defense*, *speed*, *generation* dan *legendary*. Jumlah data yang ada pada *dataset* ini adalah 800 data. Data yang ada pada *dataset* ini berdasarkan data dari game pokemon saja, tidak termasuk pokemon cards atau pokemon go.

### B. Mengidentifikasi Tipe Pokemon

*Dataset* pokemon ini mengandung sebuah kolom yang mengandung data tentang tipe pokemon. Terdapat dua tipe pokemon yang ada yaitu *type 1* dan *type 2*. Tipe pokemon ini menggambarkan kekuatan dan kelemahan pada setiap pokemon. Satu Pokemon memungkinkan memiliki lebih dari satu tipe dan maksimal adalah dua tipe. Eksplorasi ini dilakukan untuk mencari informasi tipe pokemon apa saja yang ada dalam *dataset*, selain itu membandingkan jumlah *type 1* dan *type 2*, terakhir adalah mencari kombinasi tipe pokemon yang paling banyak terbentuk.

### C. Menganalisis Tipe Pokemon Berdasarkan Beberapa Variabel

Analisa ini bertujuan untuk membandingkan tipe pokemon berdasarkan beberapa *variable*. *Variable* tersebut meliputi HP, *attack*, *defense*, *special attack*, *special defense* dan *speed*. Eksplorasi yang dilakukan pada analisa ini adalah mencari nilai *mean* beberapa *variable* yang disebutkan sebelumnya, selain itu ada juga visualisasi data berbentuk bar plot agar data yang didapat lebih mudah dipahami.

### D. Menganalisis Pokemon Legendary dan Non Legendary Berdasarkan Beberapa Variabel

Analisa ini bertujuan untuk membandingkan pokemon legendary dan non legendary berdasarkan beberapa *variable*. *Variable* tersebut meliputi HP, *attack*, *defense*, *special attack*, *special defense* dan *speed*. Perbedaan antara pokemon legendary dan non legendary adalah pokemon legendary merupakan pokemon yang kuat dan keberadaan populasi pokemon legendary ini sangat sedikit bila dibandingkan dengan pokemon non legendary. Eksplorasi yang dilakukan pada analisa ini adalah membuat visualisasi berupa *distplot*. Visualisasi ini nantinya akan menggambarkan rentang nilai data-data yang dimiliki oleh pokemon legendary dan non legendary.

### E. Menganalisis Pokemon Dari Setiap Generasi

Analisa ini bertujuan untuk mencari tahu berapa jumlah pokemon yang ada pada setiap generasinya, selain itu ada juga eksplorasi terkait jumlah legendary pokemon di setiap generasi. Analisa ini bertujuan untuk mencari tahu berapa jumlah pokemon yang ada pada setiap generasinya, selain itu ada juga eksplorasi terkait jumlah legendary pokemon di setiap generasi. Setelah melakukan kedua eksplorasi tersebut maka akan dibuatkan juga visualisasi data berupa bar plot agar data yang ditampilkan lebih mudah dilihat dan dipahami.

### F. Menganalisis Legendary Pokemon Pada Setiap Tipe

Analisa ini bertujuan untuk mencari informasi, ada berapa pokemon *legendary* yang terdapat pada setiap tipe, eksplorasi ini juga menggunakan fungsi *count()* untuk mengetahui jumlah pokemon *legendary* pada setiap tipenya. Visualisasi data juga akan sangat membantu eksplorasi ini untuk membandingkan jumlah pokemon *legendary* dan non *legendary* yang ada pada setiap tipe, visualisasi yang digunakan adalah bar plot.

### G. Menganalisis Mega Evolusi Pada Setiap Pokemon

Evolusi yang akan dibahas pada eksplorasi ini adalah Mega Evolusi. Tidak semua pokemon dapat melakukan Mega Evolusi ini. Eksplorasi ini bertujuan untuk mencari tahu pokemon apa saja yang dapat melakukan Mega Evolusi dan salah satu ciri pokemon dapat melakukan Mega Evolusi adalah mereka akan memiliki kata Mega di depan nama mereka, selain itu akan dicari juga 5 besar pokemon mega evolusi berdasarkan beberapa *variable* seperti HP, *attack* dan lainnya yang akan digambarkan dalam bentuk visualisasi bar plot.

### H. Menganalisis Perbandingan Antara Pokemon Mega Evolusi Dan Pokemon Normal

Analisa ini bertujuan untuk mencari informasi apakah pokemon yang berevolusi ini selalu lebih kuat bila dibandingkan dengan pokemon normal atau tidak. Perbandingan yang dilakukan dilandaskan oleh beberapa *variable* seperti HP, *attack*, *defense*, *special attack*, *special defense* dan *speed*. Perbandingan ini membandingkan antara pokemon Mega Evolusi terkuat dengan pokemon normal terkuat. Perbandingan ini akan dilihat dalam bentuk visualisasi *bar plot* agar perbandingan datanya mudah dilihat.

### I. Menganalisis Top 10 Pokemon Berdasarkan Beberapa Variable

Terdapat banyak sekali pokemon yang ada pada *dataset*, analisa ini bertujuan untuk mencari tahu pokemon apa saja yang memiliki data tertinggi yang didasarkan oleh beberapa *variable* seperti HP, *attack*, *defense*, *special attack*, *special defense* dan *speed*. Akan dibuatkan juga visualisasi data dengan bentuk *bar plot*.

### J. Menganalisis Perubahan Statistik Pokemon Legendary dan Non Legendary Dari Berbagai Generasi

*Dataset* yang ada mengandung informasi tentang generasi pokemon, generasi pokemon ini dimulai dari generasi pertama sampai keenam. Analisa ini menggunakan visualisasi berupa *lineplot* yang bertujuan untuk melihat perubahan nilai dari waktu ke waktu. Nilai yang dipakai adalah beberapa *variable* seperti HP, *attack*, *defense*, *special attack*, *special defense*, *speed* dan waktu yang digunakan adalah generasi pokemon.

### K. Merekomendasikan Pokemon Yang Cocok Untuk Dimiliki

Analisa ini bertujuan untuk merekomendasikan pokemon yang cocok untuk dimiliki, didasarkan oleh beberapa *variable* seperti HP, *attack*, *defense* dan *speed*. Setiap pokemon itu memiliki kelebihan dan kekurangan di setiap *variablenya*, oleh sebab itu merancang tim pokemon yang seimbang itu penting untuk melakukan pertarungan dengan pokemon lawan, dalam

komposisi tim tersebut harus ada pokemon yang memiliki kelebihan pada *variable HP, attack, defense* maupun *speed*. Beberapa *variable* tersebut nanti akan dibandingkan dengan nilai *mean* dari keseluruhan pokemon.

#### L. Menganalisis Kelemahan Pada Variable Tertentu

Sebelumnya terdapat sebuah eksplorasi yang mencari sepuluh besar pokemon dari beberapa *variable* tertentu. Analisa ini bertujuan untuk mencari informasi, apakah ada pokemon dengan keunggulan pada *variable* tertentu memiliki kelemahan pada *variable* lainnya. Data yang ada akan dibandingkan dengan nilai *mean* secara keseluruhan. Akan ada visualisasi data berupa *scatter plot*.

### IV. HASIL DAN TEMUAN

#### A. Kombinasi Antara Type 1 dan Type 2

Pada *dataset* pokemon terdapat 2 tipe yaitu tipe 1 dan tipe 2. Gambar 1 di bawah menggambarkan kombinasi yang terjadi antara tipe 1 dan tipe 2. Kombinasi yang paling banyak terbentuk adalah tipe *normal-flying* dengan jumlah sebanyak 24 dan terdapat cukup banyak yang jumlah kombinasinya itu hanya 1, seperti contoh kombinasi *bug-ghost, ground-steel, fire-rock* dan lainnya. Jumlah dari keseluruhan kombinasi yang terbentuk adalah 136.

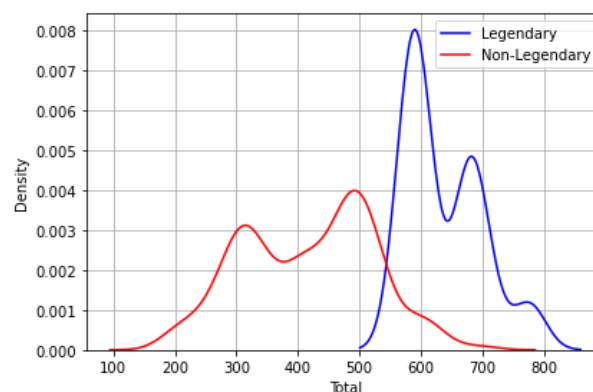
```
def type_type2_combination():  
    df['Type 1-2'] = df['Type 1'] + '-' + df['Type 2']  
    return df['Type 1-2'].value_counts()  
type_type2_combination()  
  
Normal-Flying    24  
Grass-Poison     15  
Bug-Flying       14  
Bug-Poison       12  
Water-Ground     10  
..              ..  
Electric-Grass   1  
Grass-Dragon     1  
Fire-Rock        1  
Ground-Steel     1  
Bug-Ghost        1  
Name: Type 1-2, Length: 136, dtype: int64
```

```
def type_type2_combination_number():  
    df['Type 1-2'] = df['Type 1'] + '-' + df['Type 2']  
    return df['Type 1-2'].nunique()  
type_type2_combination_number()  
  
136
```

Gambar 1. Hasil dan Jumlah Kombinasi Antara Tipe 1 dan Tipe 2

#### B. Perbandingan Pokemon Legendary dan Non Legendary Berdasarkan Total

Perbandingan antara pokemon *legendary* dan non *legendary* ini didasarkan oleh beberapa *variable* seperti *HP, attack, defense* dan lainnya. Eksplorasi di bawah merupakan hasil perbandingan antara pokemon *legendary* dan non *legendary* berdasarkan total. Total itu merepresentasikan jumlah dari *HP, attack, defense, special attack, special defense* dan *speed*. Terlihat pada Gambar 2 bahwa sebaran data pokemon non *legendary* lebih beragam dibandingkan pokemon *legendary*. Nilai total pokemon *Legendary* lebih besar dengan kisaran 500-800 sedangkan nilai total pokemon non *Legendary* lebih kecil dengan kisaran 100-700.



Gambar 2. Perbandingan Variable Total Antara Pokemon Legendary dan Non Legendary

### C. Menghitung Jumlah Pokemon Dari Setiap Generasi

Pokemon yang ada pada *dataset* ini terdiri dari berbagai generasi, terdiri dari generasi 1 sampai 6. Gambar 3 di bawah menunjukkan hasil eksplorasi menghitung jumlah pokemon di setiap generasi. Terlihat pada eksplorasi tersebut bahwa jumlah pokemon dari generasi pertama sampai keenam mengalami jumlah yang naik turun. Generasi pertama adalah generasi dengan pokemon terbanyak dengan jumlah 166 pokemon sedangkan generasi keenam adalah generasi dengan jumlah pokemon paling sedikit dengan jumlah 82 pokemon.

```
print('Analisa Jumlah Pokemon Dari Setiap Generasi,\n Jumlah dari generasi pertama : \n',df[df['Generation']==1]['Type 1'].count())
print('Jumlah dari generasi kedua : \n',df[df['Generation']==2]['Type 1'].count())
print('Jumlah dari generasi ketiga : \n',df[df['Generation']==3]['Type 1'].count())
print('Jumlah dari generasi keempat : \n',df[df['Generation']==4]['Type 1'].count())
print('Jumlah dari generasi kelima : \n',df[df['Generation']==5]['Type 1'].count())
print('Jumlah dari generasi keenam : \n',df[df['Generation']==6]['Type 1'].count())
```

```
Analisa Jumlah Pokemon Dari Setiap Generasi,
Jumlah dari generasi pertama :
166
Jumlah dari generasi kedua :
106
Jumlah dari generasi ketiga :
160
Jumlah dari generasi keempat :
121
Jumlah dari generasi kelima :
165
Jumlah dari generasi keenam :
82
```

Gambar 3. Analisa Jumlah Pokemon dari Setiap Generasi

### D. Menghitung Jumlah Legendary Pokemon Di Setiap Generasi

Sama seperti eksplorasi sebelumnya menghitung pokemon dari berbagai generasi, akan tetapi eksplorasi ini akan menghitung jumlah pokemon *legendary* dari setiap generasinya. Terlihat pada Gambar 4 di bawah generasi ketiga adalah generasi dengan jumlah *legendary* pokemon terbanyak dengan jumlah 18 pokemon *legendary* sedangkan generasi kedua adalah generasi dengan jumlah *legendary* pokemon paling sedikit dengan jumlah 5 pokemon *legendary*.

```
print('Jumlah Pokemon Legendary Pada Generasi Pertama : \n',df[(df['Legendary']==True)&(df['Generation']==1)]['Type 1'].count())
print('Jumlah Pokemon Legendary Pada Generasi Kedua : \n',df[(df['Legendary']==True)&(df['Generation']==2)]['Type 1'].count())
print('Jumlah Pokemon Legendary Pada Generasi Ketiga : \n',df[(df['Legendary']==True)&(df['Generation']==3)]['Type 1'].count())
print('Jumlah Pokemon Legendary Pada Generasi Keempat : \n',df[(df['Legendary']==True)&(df['Generation']==4)]['Type 1'].count())
print('Jumlah Pokemon Legendary Pada Generasi Kelima : \n',df[(df['Legendary']==True)&(df['Generation']==5)]['Type 1'].count())
print('Jumlah Pokemon Legendary Pada Generasi Keenam : \n',df[(df['Legendary']==True)&(df['Generation']==6)]['Type 1'].count())
```

```
Jumlah Pokemon Legendary Pada Generasi Pertama :
6
Jumlah Pokemon Legendary Pada Generasi Kedua :
5
Jumlah Pokemon Legendary Pada Generasi Ketiga :
18
Jumlah Pokemon Legendary Pada Generasi Keempat :
13
Jumlah Pokemon Legendary Pada Generasi Kelima :
15
Jumlah Pokemon Legendary Pada Generasi Keenam :
8
```

Gambar 4. Analisa Jumlah Pokemon Legendary dari Setiap Generasi

### E. Menghitung Jumlah Pokemon Yang Dapat Melakukan Mega Evolusi

Pokemon itu dapat berevolusi menjadi pokemon yang lebih kuat. Pada eksplorasi ini evolusi yang dimaksud adalah mega evolusi, tidak semua pokemon dapat melakukan mega evolusi ini. Terlihat pada Gambar 5 di bawah beberapa daftar pokemon yang dapat melakukan mega evolusi. Terdapat 48 pokemon yang dapat melakukan mega evolusi. Hal ini menunjukkan bahwa hanya sebagian kecil pokemon saja yang dapat melakukan Mega Evolusi karena jumlah pokemon yang ada pada *dataset* adalah 800 pokemon.

```

]: print('Berikut adalah beberapa pokemon yang mengalami evolusi : \n',evolusi['Name'].unique(),'\n\n')
print('Total pokemon yang mengalami evolusi adalah : \n',evolusi['Name'].nunique(), 'pokemon yang berevolusi \n\n')

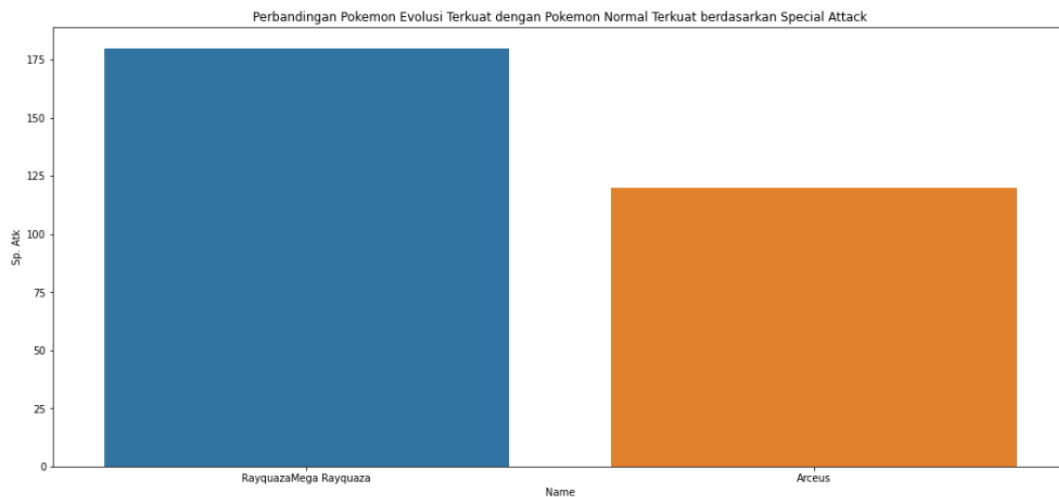
Berikut adalah beberapa pokemon yang mengalami evolusi :
['MewtwoMega Mewtwo Y' 'MewtwoMega Mewtwo X' 'RayquazaMega Rayquaza'
'DiancieMega Diancie' 'LatiasMega Latias' 'SalamenceMega Salamence'
'TyranitarMega Tyranitar' 'MetagrossMega Metagross' 'LatiosMega Latios'
'GarchompMega Garchomp' 'GyaradosMega Gyarados' 'SwampertMega Swampert'
'CharizardMega Charizard Y' 'CharizardMega Charizard X'
'AggronMega Aggron' 'BlastoiseMega Blastoise' 'BlazikenMega Blaziken'
'SceptileMega Sceptile' 'LucarioMega Lucario' 'VenusaurMega Venusaur'
'GardevoirMega Gardevoir' 'GalladeMega Gallade'
'AerodactylMega Aerodactyl' 'SteelixMega Steelix' 'AmpharosMega Ampharos'
'HeracrossMega Heracross' 'GengarMega Gengar' 'PinsirMega Pinsir'
'HoundoomMega Houndoom' 'ScizorMega Scizor' 'AbomasnowMega Abomasnow'
'SlowbroMega Slowbro' 'AltariaMega Altaria' 'KangaskhanMega Kangaskhan'
'AlakazamMega Alakazam' 'GlalieMega Glalie' 'LopunnyMega Lopunny'
'PidgeotMega Pidgeot' 'ManectricMega Manectric' 'AbsolMega Absol'
'SharpedoMega Sharpedo' 'CameruptMega Camerupt' 'BanetteMega Banette'
'AudinoMega Audino' 'MedichamMega Medicham' 'BeedrillMega Beedrill'
'MawileMega Mawile' 'SableyeMega Sableye']

Total pokemon yang mengalami evolusi adalah :
48 pokemon yang berevolusi
    
```

Gambar 5. Jumlah Pokemon yang Dapat Melakukan Mega Evolusi

#### F. Perbandingan Pokemon Mega Evolusi Dan Pokemon Normal

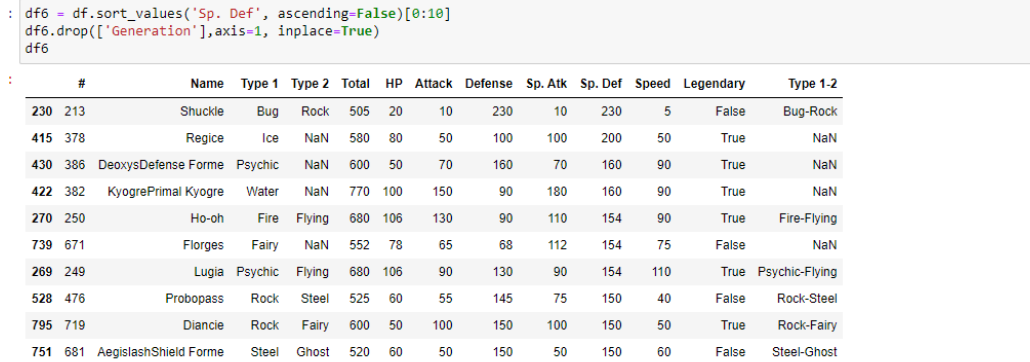
Terdapat eksplorasi terkait perbandingan antara pokemon normal terkuat dengan pokemon mega evolusi terkuat. Pokemon normal terkuat yaitu arceus, sedangkan pokemon mega evolusi terkuat adalah Rayquazamega. Setelah dilakukan eksplorasi ternyata pokemon mega evolusi itu tidak selalu unggul di setiap *variablenya*, pokemon mega evolusi ini hanya unggul di *attack* dan *special attack*, akan tetapi memang perbedaannya cukup jauh. Perbedaannya dapat dilihat pada Gambar 6 di bawah.



Gambar 6. Perbandingan Special Attack Antara Pokemon Mega Evolusi Terkuat dan Pokemon Normal Terkuat

#### G. Top 10 Pokemon Berdasarkan Defense

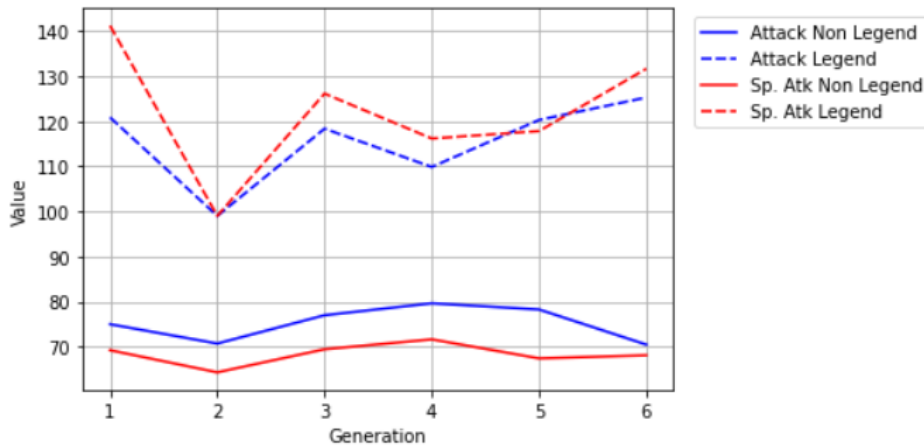
Pada *dataset* pokemon terdapat banyak sekali pokemon dengan berbagai tipe dan kemampuan yang berbeda-beda, eksplorasi ini bertujuan untuk mencari 10 pokemon berdasarkan beberapa *variable*. Pada Gambar 7 di bawah menunjukkan salah satu top 10 pokemon berdasarkan *defense*. Pada 10 pokemon dengan *defense* yang paling baik ini banyak di dominasi oleh pokemon normal. Pokemon yang memiliki *defense* terbesar yaitu pokemon Shuckle dengan *defense* sebesar 230.



Gambar 7. Top 10 Pokemon Berdasarkan Defense

H. Mencari Korelasi Perubahan Statistik Antara Attack Dan Special Attack

Eksplorasi yang dilakukan adalah membuat *lineplot* berdasarkan *variable attack* dan *special attack* dalam jangka waktu dari generasi pertama sampai keenam, dari eksplorasi ini akan dicari tahu apakah ada korelasi dari perubahan yang ada antara *attack* dan *special attack*. Terlihat pada Gambar 8 di bawah ternyata ada korelasi di antara kedua *variable* tersebut, keduanya memiliki pola garis perubahan yang hampir sama baik itu pokemon *legendary* maupun non *legendary*, hal yang perlu diketahui adalah garis yang biru itu melambangkan *variable attack* sedangkan garis yang merah melambangkan *variable special attack* dan garis putus-putus melambangkan pokemon *legendary* dan garis yang tidak putus menggambarkan pokemon non *legendary*.



Gambar 8. Korelasi Perubahan Statistik Antara Attack dan Special Attack

I. Rekomendasi Pokemon 1

Rekomendasi pokemon pertama dilandaskan pada *variable HP*, HP itu merupakan jumlah *damage* yang diterima pokemon selama pertarungan berlangsung. Ketika HP pokemon habis maka pertarungan akan berakhir, oleh sebab itu selain harus memusnahkan pokemon lawan, bertahan dari serangan lawan juga cukup penting dengan cara memilih pokemon dengan HP yang tinggi. Pokemon yang direkomendasikan dalam hal ini adalah Blissey. Bisa dilihat pada Gambar 9 di bawah data pertama adalah data *mean* dari keseluruhan pokemon, sedangkan data kedua adalah data dari pokemon blissey. Untuk HP, *special attack* dan *special defense* pokemon blissey ini berada di atas rata-rata, HP pokemon blissey ini sangat jauh di atas rata-rata. Sedangkan *attack*, *defense* dan *speed* pokemon blissey berada di bawah rata-rata. Pokemon ini bisa menjadi salah satu rekomendasi pada saat membentuk tim yang seimbang.

```

: dfHP = df.sort_values('HP', ascending=False)[0:1]
dfHP.drop(['Type 1', 'Type 2', 'Legendary', 'Type 1-2'],axis=1, inplace=True)
dfHP
:
#   Name Total HP Attack Defense Sp. Atk Sp. Def Speed Generation
261 242 Blissey 540 255 10 10 75 135 55 2

: df_outer1= pd.merge(dfnew,dfHP,how='outer')
df_outer1
C:\Users\user\anaconda3\lib\site-packages\pandas\core\reshape\merge.py:1204: UserWarning: You are merging on int and float columns where the float values are not equal to their int representation
warnings.warn(
:
#   Total HP Attack Defense Sp. Atk Sp. Def Speed Generation Name
0 362.81375 435.1025 69.25875 79.00125 73.8425 72.82 71.9025 68.2775 3.32375 NaN
1 242.00000 540.00000 255.00000 10.00000 10.0000 75.00 135.0000 55.0000 2.00000 Blissey

```

Gambar 9. Rekomendasi Pokemon 1

### J. Rekomendasi Pokemon 2

Rekomendasi pokemon kedua dilandaskan pada *variable attack*, *attack* merupakan salah satu *variable* yang penting dalam pertarungan. Memiliki susunan pokemon yang lengkap dengan *attack* tertinggi akan sangat membantu untuk memenangkan pertandingan. Pokemon yang direkomendasikan pada hal ini adalah MewtwoMega Mewtwo X. Terlihat pada Gambar 10, data pertama adalah data *mean* dari keseluruhan pokemon, sedangkan data kedua adalah data dari pokemon MewtwoMega Mewtwo X. Pokemon ini bisa dibilang merupakan salah satu pokemon yang sempurna, karena keseluruhan *variablenya* berada di atas rata-rata. Walaupun begitu, harus diketahui bahwa pokemon ini termasuk pokemon yang langka, jadi jika bisa memiliki pokemon ini, maka akan sangat membantu pada saat pertarungan melawan pokemon lawan.

```

: dfAttack = df.sort_values('Attack', ascending=False)[0:1]
dfAttack.drop(['Type 1', 'Type 2', 'Legendary', 'Type 1-2'],axis=1, inplace=True)
dfAttack
:
#   Name Total HP Attack Defense Sp. Atk Sp. Def Speed Generation
163 150 MewtwoMega Mewtwo X 780 106 190 100 154 100 130 1

: df_outer2= pd.merge(dfnew,dfAttack,how='outer')
df_outer2
C:\Users\user\anaconda3\lib\site-packages\pandas\core\reshape\merge.py:1204: UserWarning: You are merging on int and float columns where the float values are not equal to their int representation
warnings.warn(
:
#   Total HP Attack Defense Sp. Atk Sp. Def Speed Generation Name
0 362.81375 435.1025 69.25875 79.00125 73.8425 72.82 71.9025 68.2775 3.32375 NaN
1 150.00000 780.00000 106.00000 190.00000 100.0000 154.00 100.0000 130.0000 1.00000 MewtwoMega Mewtwo X

```

Gambar 10. Rekomendasi Pokemon 2

### K. Rekomendasi Pokemon 3

Rekomendasi pokemon ketiga dilandaskan pada *variable defense*, *defense* ini adalah *variable* yang menentukan seberapa besar *damage* yang diterima oleh pokemon, karena itu pokemon dengan *defense* yang tinggi akan bertahan bila menerima *attack* dari pokemon lawan atau melakukan beberapa gerakan untuk meredam *attack* yang diberikan oleh pokemon lawan. Pokemon yang direkomendasikan pada hal ini adalah Shuckle. Terlihat pada Gambar 11, data pertama adalah data *mean* dari keseluruhan pokemon, sedangkan data kedua adalah data dari pokemon Shuckle. Untuk *defense* dan *special defense* pokemon ini berada di atas rata-rata, sedangkan HP, *attack*, *special attack* dan *speed* pokemon ini berada di bawah rata-rata. Pokemon ini memiliki kelebihan pada *defense* dan *special defense* dengan nilai yang cukup jauh di atas rata-rata.



```

: dfDefense = df.sort_values('Defense', ascending=False)[0:1]
dfDefense.drop(['Type 1', 'Type 2', 'Legendary', 'Type 1-2'],axis=1, inplace=True)
dfDefense

:
#   Name Total HP Attack Defense Sp. Atk Sp. Def Speed Generation
230 213 Shuckle 505 20 10 230 10 230 5 2

: df_outer3= pd.merge(dfnew,dfDefense,how='outer')
df_outer3

C:\Users\user\anaconda3\lib\site-packages\pandas\core\reshape\merge.py:1204: UserWarning: You are merging on int and float columns where the float values are not equal to their int representation
warnings.warn(

:
#   Total HP Attack Defense Sp. Atk Sp. Def Speed Generation Name
0 362.81375 435.1025 69.25875 79.00125 73.8425 72.82 71.9025 68.2775 3.32375 NaN
1 213.00000 505.00000 20.00000 10.00000 230.00000 10.00 230.00000 5.00000 2.00000 Shuckle
    
```

Gambar 11. Rekomendasi Pokemon 3

L. Rekomendasi Pokemon 4

Rekomendasi pokemon keempat dilandaskan pada *variable speed*, *speed* ini merupakan nilai yang menentukan pokemon mana yang akan menyerang terlebih dahulu pada saat pertarungan. Pokemon yang memiliki *speed* tertinggi dari pokemon yang masuk dalam *list* pertempuran akan mengambil giliran pertama dalam pertempuran. Pokemon yang direkomendasikan dalam hal ini adalah DeoxysSpeed Forme. Bisa dilihat pada Gambar 12 di bawah data pertama adalah data *mean* dari keseluruhan pokemon, sedangkan data kedua adalah data dari pokemon DeoxysSpeed Forme. Pokemon ini cukup baik dan beberapa *variablenya* berada di atas rata-rata, hanya HP yang berada di bawah rata-rata, itu pun tidak terlalu jauh di bawah rata-rata.

```

dfSpeed = df.sort_values('Speed', ascending=False)[0:1]
dfSpeed.drop(['Type 1', 'Type 2', 'Legendary', 'Type 1-2'],axis=1, inplace=True)
dfSpeed

#   Name Total HP Attack Defense Sp. Atk Sp. Def Speed Generation
431 386 DeoxysSpeed Forme 600 50 95 90 95 90 180 3

df_outer4= pd.merge(dfnew,dfSpeed,how='outer')
df_outer4

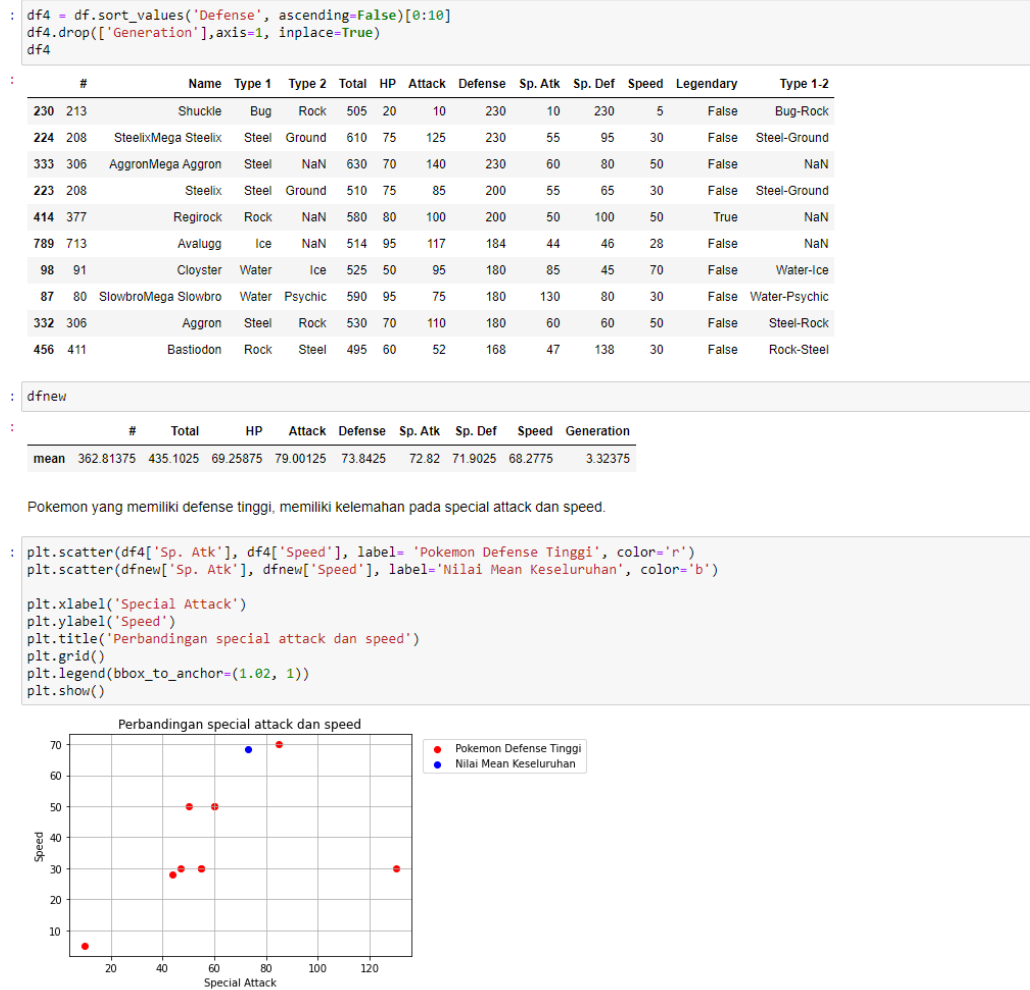
C:\Users\user\anaconda3\lib\site-packages\pandas\core\reshape\merge.py:1204: UserWarning: You are merging on int and float columns where the float values are not equal to their int representation
warnings.warn(

#   Total HP Attack Defense Sp. Atk Sp. Def Speed Generation Name
0 362.81375 435.1025 69.25875 79.00125 73.8425 72.82 71.9025 68.2775 3.32375 NaN
1 386.00000 600.00000 50.00000 95.00000 90.00000 95.00 90.00000 180.00000 3.00000 DeoxysSpeed Forme
    
```

Gambar 12. Rekomendasi Pokemon 4

M. Kelemahan Pokemon Dengan Defense yang Tinggi

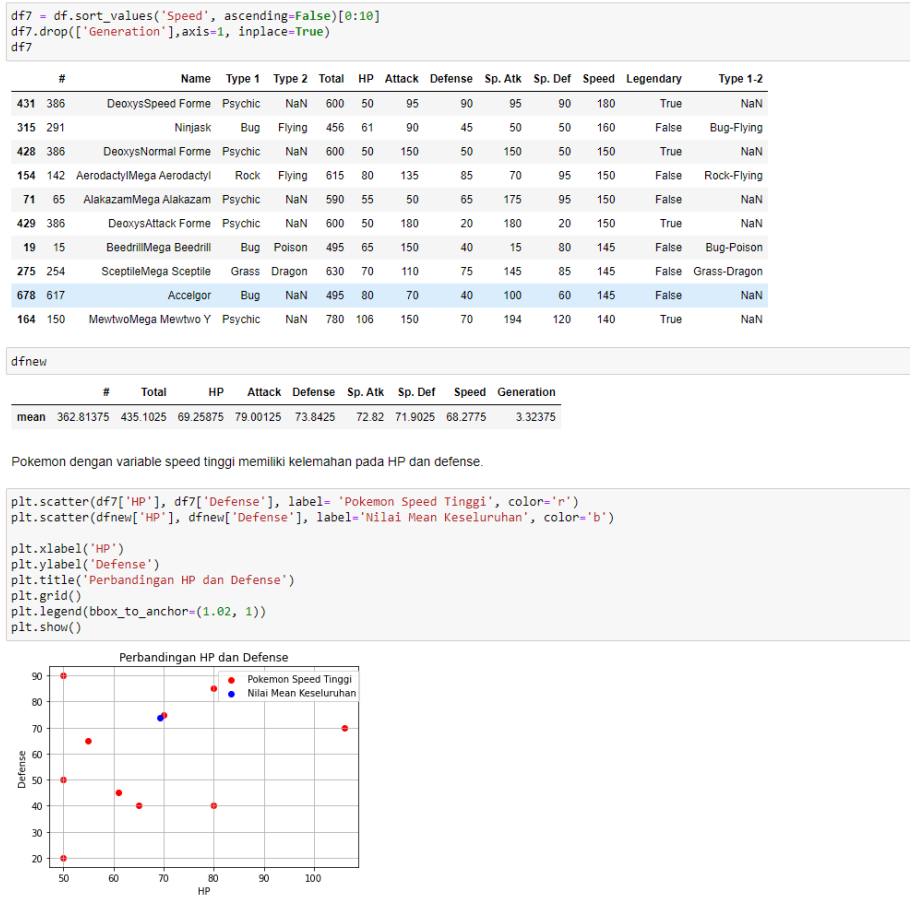
Pokemon yang memiliki *defense* tinggi, memiliki kelemahan pada *special attack* dan *speed*. Hal ini bisa dilihat pada perbandingan nilai *mean* dan nilai pokemon yang ada, merujuk pada kolom *special attack* dan *speed*. Hasil eksplorasi ini bisa dilihat pada Gambar 13 di bawah.



Gambar 13. Kelemahan Pokemon Dengan Defense Tinggi

*N. Kelemahan Pokemon Dengan Speed yang Tinggi*

Pokemon dengan *variable speed* tinggi memiliki kelemahan pada HP dan *defense*. Hal ini bisa dilihat pada perbandingan nilai *mean* dan nilai pokemon yang ada, merujuk pada kolom HP dan *defense*. Hasil eksplorasi ini bisa dilihat pada Gambar 14 di bawah.



Gambar 14. Kelemahan Pokemon Dengan Speed Tinggi

### V. KESIMPULAN

Dataset pokemon ini terdiri dari berbagai macam data yang meliputi nama pokemon, *type 1*, *type 2*, *HP*, *attack*, *defense*, *special attack*, *special defense*, *speed*, *legendary* dan *generation*. Dari data yang ada dilakukan beberapa analisa untuk menghasilkan informasi yang lebih berguna. Informasi ini dapat bermanfaat untuk pemain pokemon dalam menentukan pokemon apa yang akan mereka pilih atau miliki. Contoh informasi yang berguna itu adalah hasil analisa mengenai beberapa pokemon yang memiliki HP terbesar atau pokemon yang memiliki *attack* terbesar. Terdapat juga beberapa informasi yang belum diketahui sebelumnya yaitu pokemon dengan statistik *defense* tinggi memiliki kelemahan pada *special attack* dan *speed*, sedangkan pokemon yang memiliki statistik *speed* tinggi memiliki kelemahan pada HP dan *defense*, pokemon *legendary* populasinya lebih sedikit bila dibandingkan dengan pokemon non *legendary*. Informasi tersebut dapat digunakan untuk merancang tim pokemon yang seimbang.

Visualisasi data bertujuan untuk data yang sudah diolah dapat di representasikan dalam bentuk gambar agar lebih mudah dipahami, oleh sebab itu dari *dataset* pokemon yang sudah di analisa dan diolah, dilakukanlah visualisasi data. Berbagai macam visualisasi data yang digunakan dalam eksplorasi diantaranya adalah *bar plot*, *line plot* dan *dist plot*.

### DAFTAR PUSTAKA

- [1] K. F. Punch, Introduction to Social Research : Quantitative and Qualitative Approaches, California: SAGE Publication Inc, 2014.
- [2] M. Heydt, Learning Pandas, Birmingham: Packt Publishing Ltd, 2017.
- [3] F. Nelli, Python Data Analytics : Data Analysis and Science Usintg Pandas, Matplotlib and The Python Programming

Language, New York: Apress, 2015.

- [4] A. Devert, *Matplotlib Plotting Cookbook*, Birmingham: Packt Publishing Ltd, 2014.
- [5] A. Yim, C. Chung and A. Yu, *Matplotlib for Python Developers : Effective Techniques for Data Visualization with Python*, Birmingham: Packt Publishing Ltd, 2018.
- [6] B. W. Keller, *Mastering Matplotlib 2x : Effective Data Visualization Techniques with Python*, Birmingham: Packt Publishing Ltd, 2018.
- [7] U. M. Cakmak and M. Cuhadaroglu, *Mastering Numerical Computing with Numpy : Master Scientific Computing and Perform Complex Operations with Ease*, Birmingham: Packt Publishing Ltd, 2018.
- [8] I. Idris, *Numpy : Beginner's Guide*, Birmingham: Packt Publishing Ltd, 2015.
- [9] D. Toomey, *Learning Jupyter*, Birmingham: Packt Publishing Ltd, 2016.
- [10] W. L. Martinez, A. R. Martinez and J. L. Solka, *Exploratory Data Analysis*, Boca Raton: Taylor & Francis Group, 2017.
- [11] C. Panumate, S. Xiong, H. Lida and T. Kondo, *Evolutionary Changes of Pokemon Game : A Case Study with Focus on Catching Pokemon*, p.2, 2015.