MDR-FirstOccurrence

February 25, 2020

[1]:
```
# Imports
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import matplotlib.ticker as ticker
```

[18]:
```
def report_first_occurence(drug, drug_res):
    # Build geno dict for text replacement
    geno_db_df = pd.read_excel('geno-database.xlsx')
    ## converting df to series, and then to dict
    geno_dict = geno_db_df.set_index('ID')['Shortname']
    # Create Data Frame
    df = pd.read_csv(drug, sep='\t')
    # Text replacement
    df['to'] = df['to'].replace(geno_dict)
    df['from'] = df['from'].replace(geno_dict)
    # Apply Filter to select MDR geno
    # And find first such mutation
    frn = df[df['to'].str.contains(drug_res)].index[0]
    # Get which year that is and function returns
    return (df.loc[frn].iloc[0] / 365)
```

1 Main Function

1.1 AL

[19]:
```
report_first_occurence('fo-al.txt', '^(TYF|TNY|TNF|KYY|KYF|KNY|KNF)..Y..$')
```

```
/Users/zhewenli/anaconda3/lib/python3.7/site-packages/ipykernel_launcher.py:13:
UserWarning: This pattern has match groups. To actually get the groups, use
str.extract.
  del sys.path[0]
```

[19]:
```
10.008219178082191
```

1.2 AS-AQ

[20]:
```
report_first_occurence('fo-asaq.txt', '^(TYF|TNY|TNF|KYY|KYF|KNY|TYY)..Y..$')
```

```
/Users/zhewenli/anaconda3/lib/python3.7/site-packages/ipykernel_launcher.py:13:
UserWarning: This pattern has match groups. To actually get the groups, use
str.extract.
  del sys.path[0]
```

[20]:
```
10.008219178082191
```

1.3 DHA-PPQ

[21]:
```
report_first_occurence('fo-dhappq.txt', '^.....Y2.$')
```

[21]:
```
10.180821917808219
```

[ ]:
```

```