Result of script

Checking data

race gender age admission\_type\_id ... max\_glu\_serum A1Cresult diabetesMed readmitted

0 Caucasian Female [50-60) Elective ... None None No No

1 Caucasian Female [20-30) Urgent ... None None No No

2 Caucasian Male [80-90) Not Available ... >200 None Yes Yes

3 AfricanAmerican Female [50-60) Emergency ... None None Yes No

4 AfricanAmerican Female [50-60) Emergency ... None None Yes No

[5 rows x 21 columns]

Part 1:

Total NaN count for each column:

race 0

gender 0

age 0

admission\_type\_id 1374

discharge\_disposition\_id 604

admission\_source\_id 973

time\_in\_hospital 0

num\_lab\_procedures 0

num\_procedures 0

num\_medications 0

number\_outpatient 0

number\_emergency 0

number\_inpatient 0

diag\_1 2

diag\_2 59

diag\_3 209

number\_diagnoses 1

max\_glu\_serum 9336

A1Cresult 8379

diabetesMed 0

readmitted 0

Percentage of NaN values:

1 : 0.0 %

2 : 0.0 %

3 : 0.0 %

4 : 13.74 %

5 : 6.04 %

6 : 9.73 %

7 : 0.0 %

8 : 0.0 %

9 : 0.0 %

10 : 0.0 %

11 : 0.0 %

12 : 0.0 %

13 : 0.0 %

14 : 0.02 %

15 : 0.59 %

16 : 2.09 %

17 : 0.01 %

18 : 93.36 %

19 : 83.78999999999999 %

20 : 0.0 %

21 : 0.0 %

race gender age admission\_type\_id ... diag\_3 number\_diagnoses diabetesMed readmitted

0 Caucasian Female [50-60) Elective ... 719 9.0 No No

1 Caucasian Female [20-30) Urgent ... 285 6.0 No No

2 Caucasian Male [80-90) Emergency ... 276 9.0 Yes Yes

3 AfricanAmerican Female [50-60) Emergency ... 250 3.0 Yes No

4 AfricanAmerican Female [50-60) Emergency ... 298 7.0 Yes No

... ... ... ... ... ... ... ... ... ...

9995 Caucasian Female [60-70) Elective ... 272 7.0 Yes No

9996 AfricanAmerican Male [60-70) Urgent ... V53 9.0 Yes No

9997 AfricanAmerican Female [70-80) Emergency ... 536 6.0 No Yes

9998 Caucasian Male [80-90) Urgent ... 599 9.0 No No

9999 Caucasian Male [70-80) Elective ... 427 9.0 Yes No

[9646 rows x 19 columns]

Total NaN count for each column:

race 0

gender 0

age 0

admission\_type\_id 0

discharge\_disposition\_id 0

admission\_source\_id 0

time\_in\_hospital 0

num\_lab\_procedures 0

num\_procedures 0

num\_medications 0

number\_outpatient 0

number\_emergency 0

number\_inpatient 0

diag\_1 2

diag\_2 58

diag\_3 206

number\_diagnoses 0

diabetesMed 0

readmitted 0

Percentage of NaN values:

1 : 0.0 %

2 : 0.0 %

3 : 0.0 %

4 : 0.0 %

5 : 0.0 %

6 : 0.0 %

7 : 0.0 %

8 : 0.0 %

9 : 0.0 %

10 : 0.0 %

11 : 0.0 %

12 : 0.0 %

13 : 0.0 %

14 : 0.02 %

15 : 0.58 %

16 : 2.06 %

17 : 0.0 %

18 : 0.0 %

19 : 0.0 %

Exploring Analytics - Part 2:

readmitted - No

5897

readmitted - Yes

3749

diabetesMed - No

2457

diabetesMed - Yes

7189

discharge\_disposition\_id - Expired

186

discharge\_disposition\_id - Not Expired

9460

Part 3:

race\_AfricanAmerican race\_Asian race\_Caucasian race\_Hispanic ... diag\_3\_V72 diag\_3\_V85 readmitted\_No readmitted\_Yes

0 0 0 1 0 ... 0 0 1 0

1 0 0 1 0 ... 0 0 1 0

2 0 0 1 0 ... 0 0 0 1

3 1 0 0 0 ... 0 0 1 0

4 1 0 0 0 ... 0 0 1 0

... ... ... ... ... ... ... ... ... ...

9995 0 0 1 0 ... 0 0 1 0

9996 1 0 0 0 ... 0 0 1 0

9997 1 0 0 0 ... 0 0 0 1

9998 0 0 1 0 ... 0 0 1 0

9999 0 0 1 0 ... 0 0 1 0

[9646 rows x 1368 columns]

Processing 48 combinations | Sampling itemset size 43

support itemsets

0 0.733257 (race\_Caucasian)

1 0.540950 (gender\_Female)

2 0.459050 (gender\_Male)

3 0.626063 (admission\_type\_id\_Emergency)

4 0.980717 (discharge\_disposition\_id\_Not Expired)

5 0.588949 (admission\_source\_id\_Emergency Room)

6 0.302509 (admission\_source\_id\_Physician Referral)

7 0.611341 (readmitted\_No)

8 0.388659 (readmitted\_Yes)

9 0.381920 (gender\_Female, race\_Caucasian)

10 0.351337 (gender\_Male, race\_Caucasian)

11 0.450549 (admission\_type\_id\_Emergency, race\_Caucasian)

12 0.718536 (discharge\_disposition\_id\_Not Expired, race\_Ca...

13 0.432200 (admission\_source\_id\_Emergency Room, race\_Cauc...

14 0.434584 (readmitted\_No, race\_Caucasian)

15 0.343044 (admission\_type\_id\_Emergency, gender\_Female)

16 0.530997 (gender\_Female, discharge\_disposition\_id\_Not E...

17 0.317956 (admission\_source\_id\_Emergency Room, gender\_Fe...

18 0.327286 (readmitted\_No, gender\_Female)

19 0.449720 (gender\_Male, discharge\_disposition\_id\_Not Exp...

20 0.611860 (admission\_type\_id\_Emergency, discharge\_dispos...

21 0.524570 (admission\_source\_id\_Emergency Room, admission...

22 0.376114 (admission\_type\_id\_Emergency, readmitted\_No)

23 0.575783 (admission\_source\_id\_Emergency Room, discharge...

24 0.592059 (readmitted\_No, discharge\_disposition\_id\_Not E...

25 0.388659 (readmitted\_Yes, discharge\_disposition\_id\_Not ...

26 0.342940 (admission\_source\_id\_Emergency Room, readmitte...

27 0.374456 (gender\_Female, discharge\_disposition\_id\_Not E...

28 0.344080 (discharge\_disposition\_id\_Not Expired, gender\_...

29 0.439768 (admission\_type\_id\_Emergency, discharge\_dispos...

30 0.379950 (admission\_source\_id\_Emergency Room, admission...

31 0.421833 (admission\_source\_id\_Emergency Room, discharge...

32 0.419863 (discharge\_disposition\_id\_Not Expired, readmit...

33 0.335683 (admission\_type\_id\_Emergency, gender\_Female, d...

34 0.311217 (admission\_source\_id\_Emergency Room, gender\_Fe...

35 0.317334 (gender\_Female, readmitted\_No, discharge\_dispo...

36 0.512233 (admission\_source\_id\_Emergency Room, admission...

37 0.361912 (admission\_type\_id\_Emergency, readmitted\_No, d...

38 0.309040 (admission\_source\_id\_Emergency Room, admission...

39 0.329774 (admission\_source\_id\_Emergency Room, readmitte...

40 0.370309 (admission\_source\_id\_Emergency Room, admission...

All Rules:

antecedents consequents antecedent support ... lift leverage conviction

0 (gender\_Female) (race\_Caucasian) 0.540950 ... 0.962851 -0.014735 0.907342

1 (race\_Caucasian) (gender\_Female) 0.733257 ... 0.962851 -0.014735 0.958059

2 (gender\_Male) (race\_Caucasian) 0.459050 ... 1.043777 0.014735 1.136801

3 (race\_Caucasian) (gender\_Male) 0.733257 ... 1.043777 0.014735 1.038582

4 (admission\_type\_id\_Emergency) (race\_Caucasian) 0.626063 ... 0.981450 -0.008516 0.951482

[5 rows x 9 columns]

Part 4:

C:\Users\14235\AppData\Roaming\Python\Python39\site-packages\pandas\core\frame.py:3641: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy

self[k1] = value[k2]

c:\Users\14235\Desktop\Hospital Readmission Project\Data\_Processing.py:73: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy

dataFrame['readmitted'] = dataFrame['readmitted'].replace('Yes', 1)

c:\Users\14235\Desktop\Hospital Readmission Project\Data\_Processing.py:74: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy

dataFrame['readmitted'] = dataFrame['readmitted'].replace('No', 0)

c:\Users\14235\Desktop\Hospital Readmission Project\Data\_Processing.py:75: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy

dataFrame['diabetesMed'] = dataFrame['diabetesMed'].replace('Yes', 1)

c:\Users\14235\Desktop\Hospital Readmission Project\Data\_Processing.py:76: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy

dataFrame['diabetesMed'] = dataFrame['diabetesMed'].replace('No', 0)

c:\Users\14235\Desktop\Hospital Readmission Project\Data\_Processing.py:77: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy

dataFrame['gender'] = dataFrame['gender'].replace('Male', 1)

c:\Users\14235\Desktop\Hospital Readmission Project\Data\_Processing.py:78: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy

dataFrame['gender'] = dataFrame['gender'].replace('Female', 0)

c:\Users\14235\Desktop\Hospital Readmission Project\Data\_Processing.py:85: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#returning-a-view-versus-a-copy

dataFrame[col] = st\_data[col]

c:\Users\14235\Desktop\Hospital Readmission Project\Data\_Processing.py:119: PerformanceWarning: DataFrame is highly fragmented. This is usually the result of calling `frame.insert` many times, which has poor performance. Consider joining all columns at once using pd.concat(axis=1) instead. To get a de-fragmented frame, use `newframe = frame.copy()`

dataFrame[col] = st\_data[col]

c:\Users\14235\Desktop\Hospital Readmission Project\Data\_Processing.py:125: PerformanceWarning: DataFrame is highly fragmented. This is usually the result of calling `frame.insert` many times, which has poor performance. Consider joining all columns at once using pd.concat(axis=1) instead. To get a de-fragmented frame, use `newframe = frame.copy()`

dataFrame[col] = st\_data[col]

Decision Tree:

Confusion Matrix:

[[888 295]

[534 213]]

Accuracy:

precision recall f1-score support

0 0.62 0.75 0.68 1183

1 0.42 0.29 0.34 747

accuracy 0.57 1930

macro avg 0.52 0.52 0.51 1930

weighted avg 0.55 0.57 0.55 1930

Gaussian Naive Bayes:

Confusion Matrix:

[[ 107 1076]

[ 26 721]]

Accuracy:

precision recall f1-score support

0 0.80 0.09 0.16 1183

1 0.40 0.97 0.57 747

accuracy 0.43 1930

macro avg 0.60 0.53 0.36 1930

weighted avg 0.65 0.43 0.32 1930