In [3]:
```
#Import Libraries
import pandas as pd
```

In [4]:
```
MainData = pd.read_csv('diabetes_data.csv')
```

In [5]:
```
MainData.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 768 entries, 0 to 767
Data columns (total 9 columns):
 #   Column                    Non-Null Count  Dtype  
---  ------                    --------------  -----  
 0   Pregnancies               768 non-null    int64  
 1   Glucose                   768 non-null    int64  
 2   BloodPressure             768 non-null    int64  
 3   SkinThickness             768 non-null    int64  
 4   Insulin                   768 non-null    int64  
 5   BMI                       768 non-null    float64
 6   DiabetesPedigreeFunction  768 non-null    float64
 7   Age                       768 non-null    int64  
 8   Outcome                   768 non-null    object 
dtypes: float64(2), int64(6), object(1)
memory usage: 54.1+ KB
```

In [6]:
```
MainData.columns
```

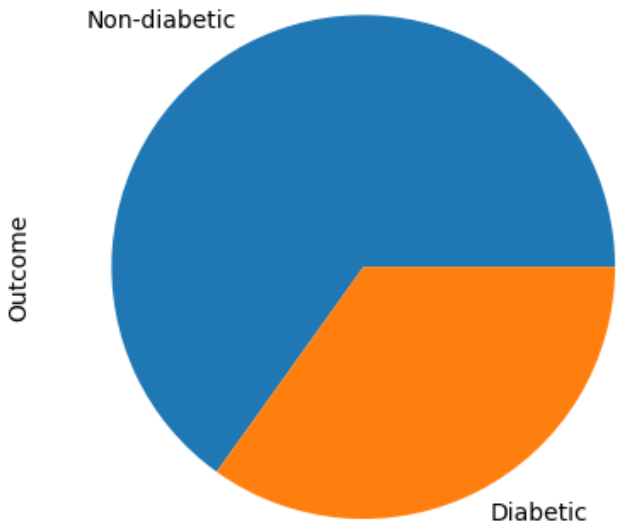
Out[6]:
```
Index(['Pregnancies', 'Glucose', 'BloodPressure', 'SkinThickness', 'Insulin',
       'BMI', 'DiabetesPedigreeFunction', 'Age', 'Outcome'],
      dtype='object')
```

In [7]:
```
#cardinality - Number of unique value
MainData.nunique()
```

Out[7]:
```
Pregnancies                17
Glucose                   136
BloodPressure              47
SkinThickness              51
Insulin                   186
BMI                       248
DiabetesPedigreeFunction  517
Age                        52
Outcome                     2
dtype: int64
```

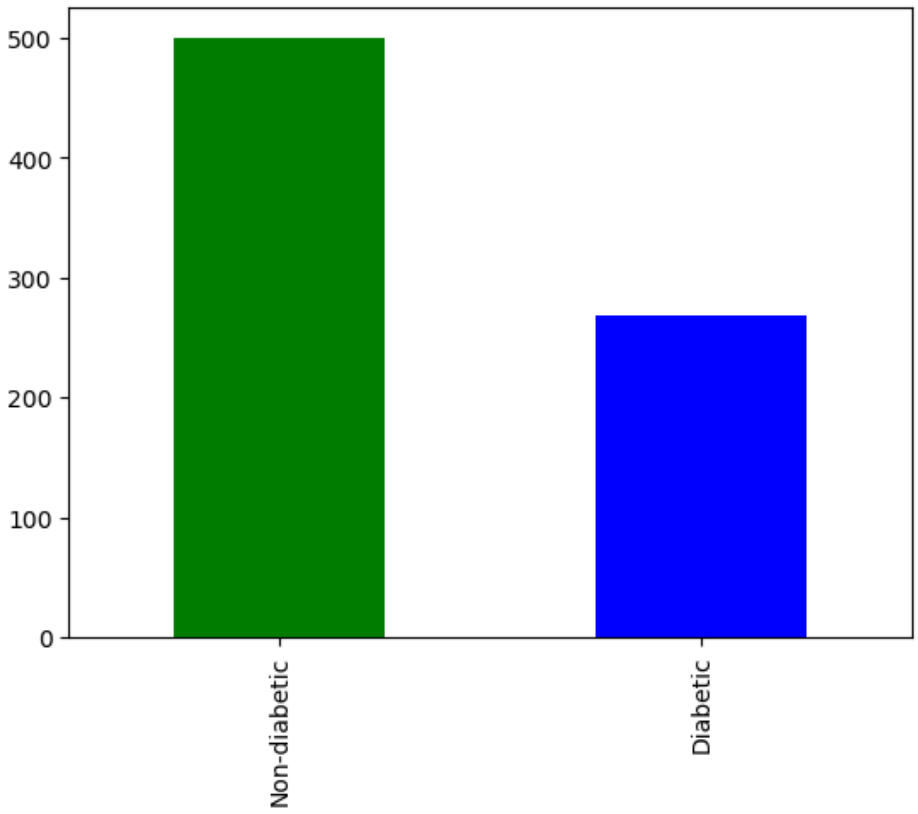
In [14]:
```
MainData['Outcome'].value_counts().plot.pie()
```

Out[14]:
```
<AxesSubplot:ylabel='Outcome'>
```



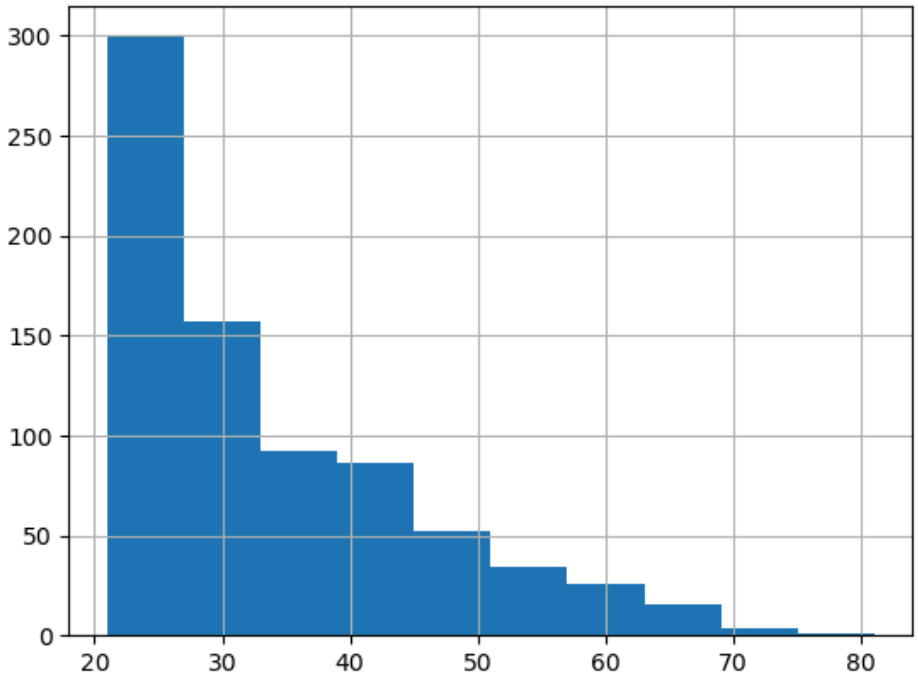
In [16]:
```
MainData['Outcome'].value_counts().plot.bar(color =['g','b'])
```

Out[16]: <AxesSubplot:>



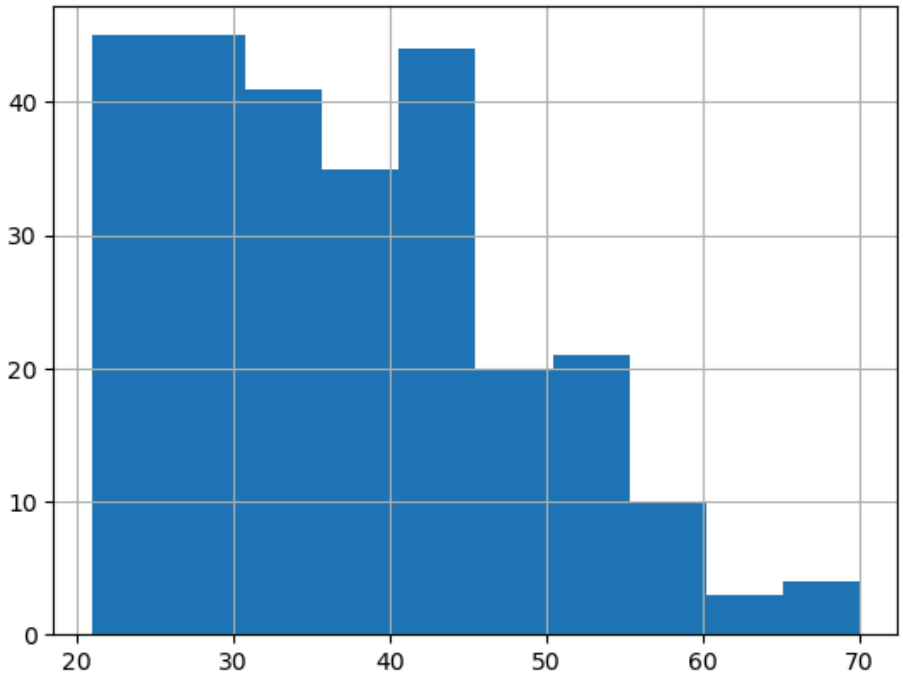
In [18]:
```
MainData['Age'].hist()
```

Out[18]: <AxesSubplot:>



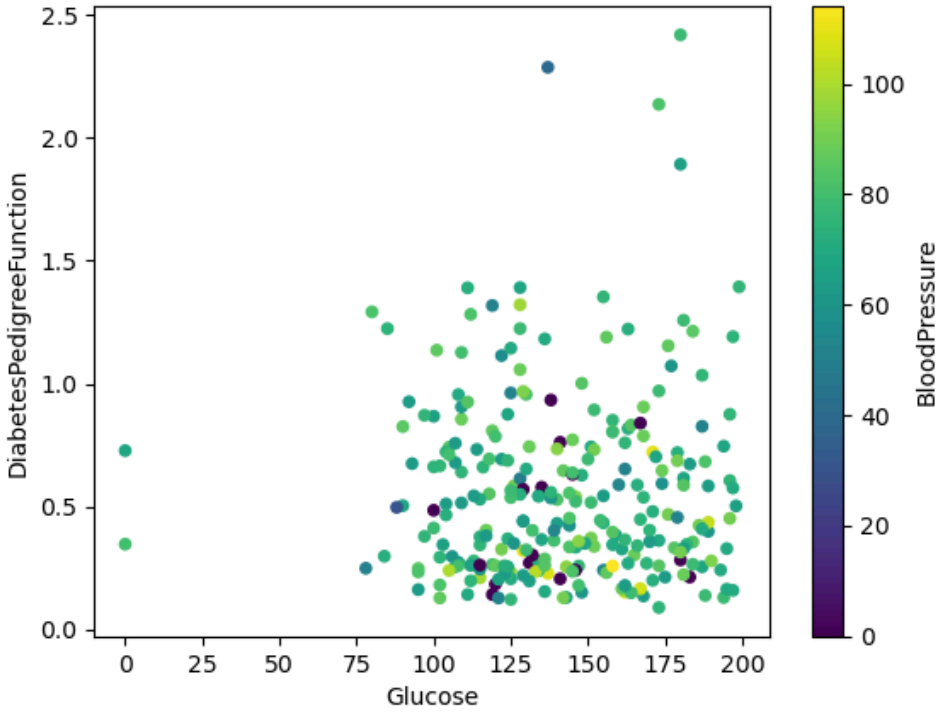
In [28]:
```
#filtering Diabetic Only
Pos_Dia = MainData[MainData['Outcome']=='Diabetic']
#visualization by age
Pos_Dia['Age'].hist()
```

Out[28]:
```
<AxesSubplot:>
```



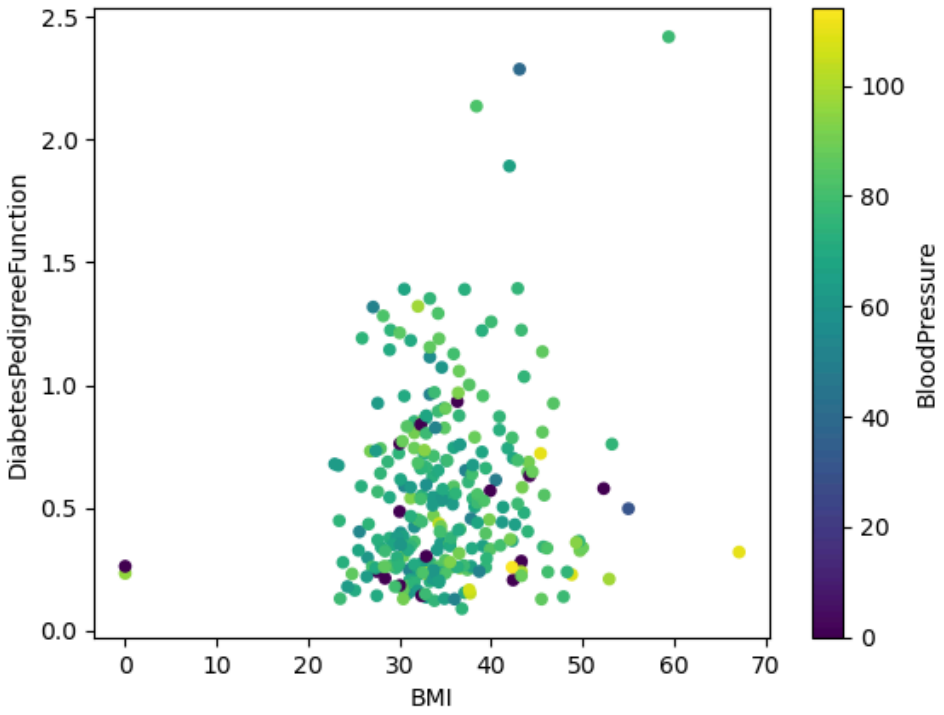
In [30]:
```
Pos_Dia.plot.scatter(x='Glucose', y='DiabetesPedigreeFunction', c='BloodPressure', colormap='viridis')
```

Out[30]:
```
<AxesSubplot:xlabel='Glucose', ylabel='DiabetesPedigreeFunction'>
```



In [32]:
```
#Pos_Dia.plot.scatter(x='BMI', y='DiabetesPedigreeFunction', s ='Glucose' , c='BloodPressure', colormap=
Pos_Dia.plot.scatter(x='BMI', y='DiabetesPedigreeFunction', c='BloodPressure', colormap='viridis')
```

Out[32]:
```
<AxesSubplot:xlabel='BMI', ylabel='DiabetesPedigreeFunction'>
```



In [ ]:
```

```