In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
```

In [2]:
```
bihar=pd.read_csv('Bihar_sample_data.csv')
bihar
```

Out[2]:

	personid	female	adult	age	height_cm	weight_kg
0	11010101	0.0	1.0	70.0	164.1	48.9
1	11010102	0.0	1.0	32.0	157.3	44.0
2	11010103	1.0	1.0	28.0	150.3	37.7
3	11010104	0.0	0.0	12.0	146.2	30.7
4	11010105	1.0	0.0	11.0	135.1	30.2
...	...	...	...	...	...	...
39548	24282403	0.0	1.0	38.0	156.1	60.7
39549	24282404	1.0	1.0	28.0	148.9	46.7
39550	24282405	1.0	0.0	7.0	116.3	17.7
39551	24282406	0.0	0.0	6.0	101.5	13.4
39552	24282407	1.0	0.0	3.0	91.5	12.2

39553 rows × 6 columns

In [3]:
```
bihar.dropna(inplace=True)
```

In [4]:
```
bihar.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 33001 entries, 0 to 39552
Data columns (total 6 columns):
 #   Column     Non-Null Count  Dtype  
---  ------     --------------  -----  
 0   personid   33001 non-null  int64  
 1   female     33001 non-null  float64
 2   adult      33001 non-null  float64
 3   age        33001 non-null  float64
 4   height_cm  33001 non-null  float64
 5   weight_kg  33001 non-null  float64
dtypes: float64(5), int64(1)
memory usage: 1.8 MB
```

In [5]:
```
bihar.isnull().sum()
```

Out[5]:
```
personid     0
female       0
adult        0
age          0
height_cm    0
weight_kg    0
dtype: int64
```

In [6]:
```
bihar['female']=bihar['female'].astype(int)
bihar['adult']=bihar['adult'].astype(int)
bihar['age']=bihar['age'].astype(int)
bihar['height_cm']=bihar['height_cm'].astype(int)
bihar['weight_kg']=bihar['weight_kg'].astype(int)
```

In [7]:
```
bihar
```

Out[7]:

	personid	female	adult	age	height_cm	weight_kg
0	11010101	0	1	70	164	48
1	11010102	0	1	32	157	44
2	11010103	1	1	28	150	37
3	11010104	0	0	12	146	30
4	11010105	1	0	11	135	30
...	...	...	...	...	...	...
39548	24282403	0	1	38	156	60
39549	24282404	1	1	28	148	46
39550	24282405	1	0	7	116	17
39551	24282406	0	0	6	101	13
39552	24282407	1	0	3	91	12

33001 rows × 6 columns

In [8]:
```
# data manipulating the data
```

In [9]:
```
bihar.head()
```

Out[9]:

	personid	female	adult	age	height_cm	weight_kg
0	11010101	0	1	70	164	48
1	11010102	0	1	32	157	44
2	11010103	1	1	28	150	37
3	11010104	0	0	12	146	30
4	11010105	1	0	11	135	30

In [10]:
```
bihar.tail()
```

Out[10]:

	personid	female	adult	age	height_cm	weight_kg
39548	24282403	0	1	38	156	60
39549	24282404	1	1	28	148	46
39550	24282405	1	0	7	116	17
39551	24282406	0	0	6	101	13
39552	24282407	1	0	3	91	12

In [11]:
```
bihar.sample(frac=0.5).shape
```

Out[11]: (16500, 6)

In [12]:
```
bihar.sample(8)
```

Out[12]:

	personid	female	adult	age	height_cm	weight_kg
28958	21100809	0	0	9	129	24
3236	12051006	1	0	8	140	30
21217	18161705	0	0	10	132	26
35954	23240805	1	0	5	109	15
25382	20011005	0	1	18	163	55
16753	16270412	1	1	23	150	48
9848	14142010	0	0	4	108	15
5553	12280509	1	0	5	101	14

In [13]:
```
# Sumarry the Statistic anlysis
```

In [14]:
```
bihar.describe()
```

Out[14]:

	personid	female	adult	age	height_cm	weight_kg
count	3.300100e+04	33001.000000	33001.000000	33001.000000	33001.000000	33001.000000
mean	1.767654e+07	0.518439	0.571195	27.580831	142.386564	39.019424
std	4.061337e+06	0.499667	0.494913	20.789065	23.608276	16.469070
min	1.101010e+07	0.000000	0.000000	0.000000	1.000000	1.000000
25%	1.414201e+07	0.000000	0.000000	10.000000	133.000000	25.000000
50%	1.802080e+07	1.000000	1.000000	22.000000	149.000000	41.000000
75%	2.118030e+07	1.000000	1.000000	40.000000	159.000000	50.000000
max	2.428241e+07	1.000000	1.000000	110.000000	195.000000	141.000000

In [15]:
```
bihar.median()
```

Out[15]:
```
personid     18020804.0
female              1.0
adult               1.0
age                22.0
height_cm         149.0
weight_kg          41.0
dtype: float64
```

In [16]:
```
bihar.mode().iloc[8]
```

Out[16]:
```
personid     11010204.0
female              NaN
adult               NaN
age                 NaN
height_cm           NaN
weight_kg           NaN
Name: 8, dtype: float64
```

In [17]:
```
bihar.isnull().sum()
```

Out[17]:
```
personid     0
female       0
adult        0
age          0
height_cm    0
weight_kg    0
dtype: int64
```

In [18]:
```
bihar['adult'].fillna(bihar['adult'].mode()[0],inplace=True)
```

In [19]:
```
bihar.isnull().sum()
```

Out[19]:
```
personid     0
female       0
adult        0
age          0
height_cm    0
weight_kg    0
dtype: int64
```

Data Visualize

In [20]:
```
#plotting the Data
```

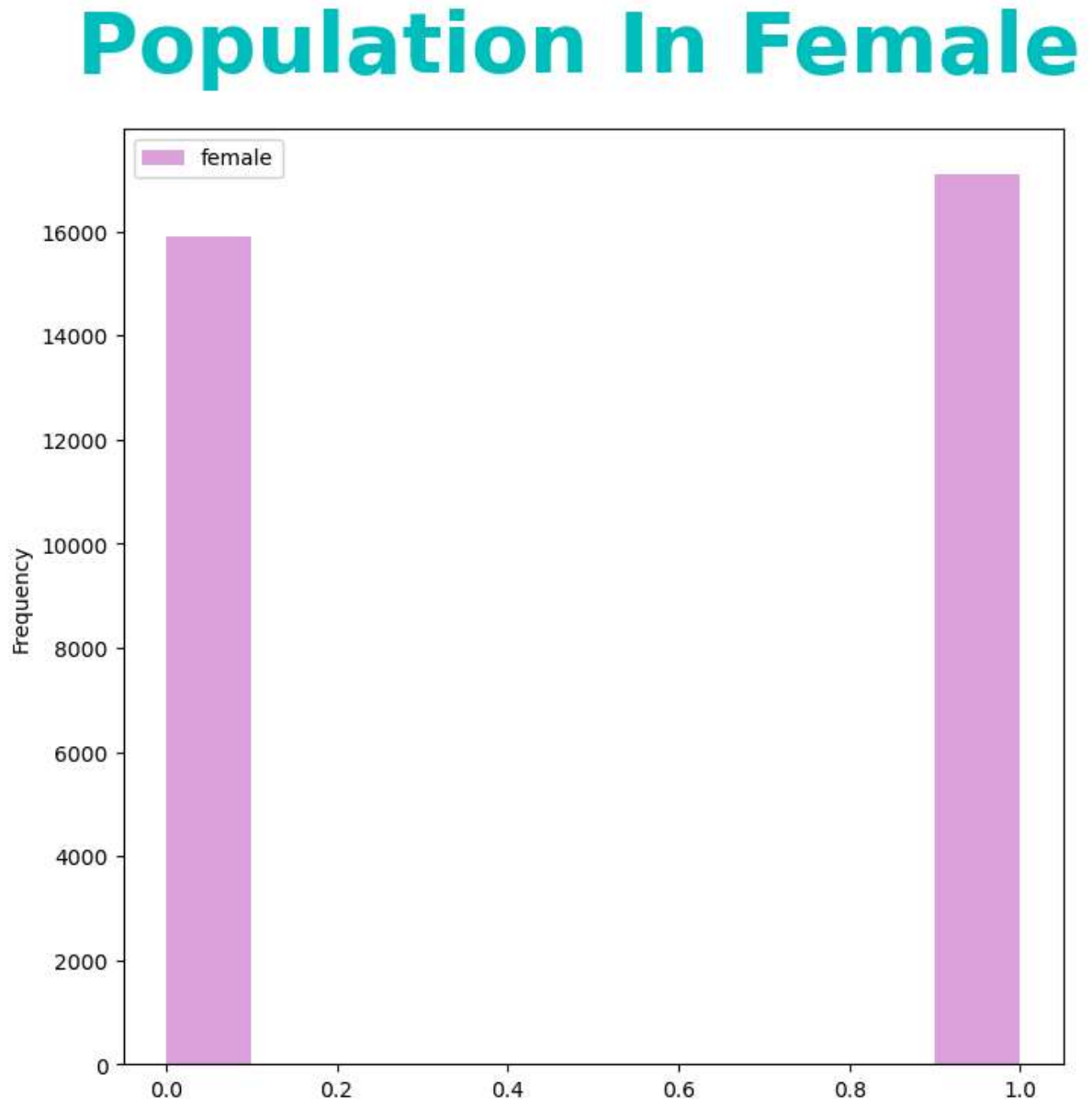
In [21]:
```
bihar.head(2)
```

Out[21]:

	personid	female	adult	age	height_cm	weight_kg
0	11010101	0	1	70	164	48
1	11010102	0	1	32	157	44

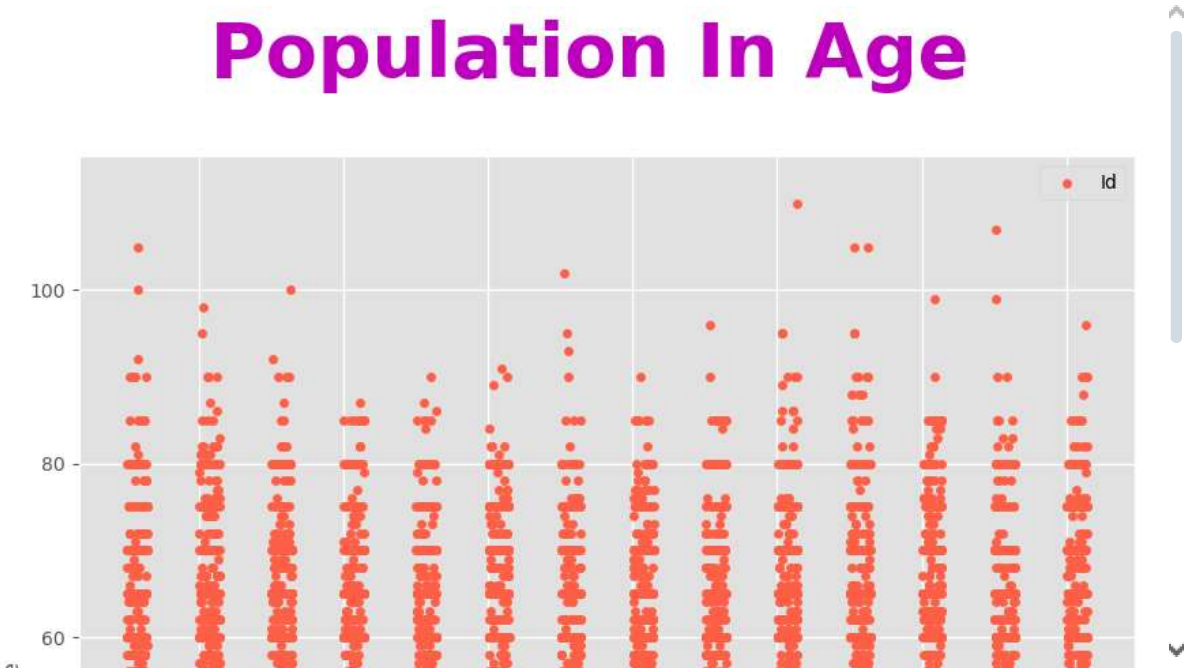
In [22]:
```
# plotting the data in  person id and female
```

In [23]:
```
bihar.plot(x='personid',
           y='female',
           kind="hist",
           color='plum',
            figsize=(8,8),
           bins=10);
plt.style.use('ggplot');
plt.suptitle('Population In Female',
              fontsize=40,
              fontweight='bold',
              color='c');
```



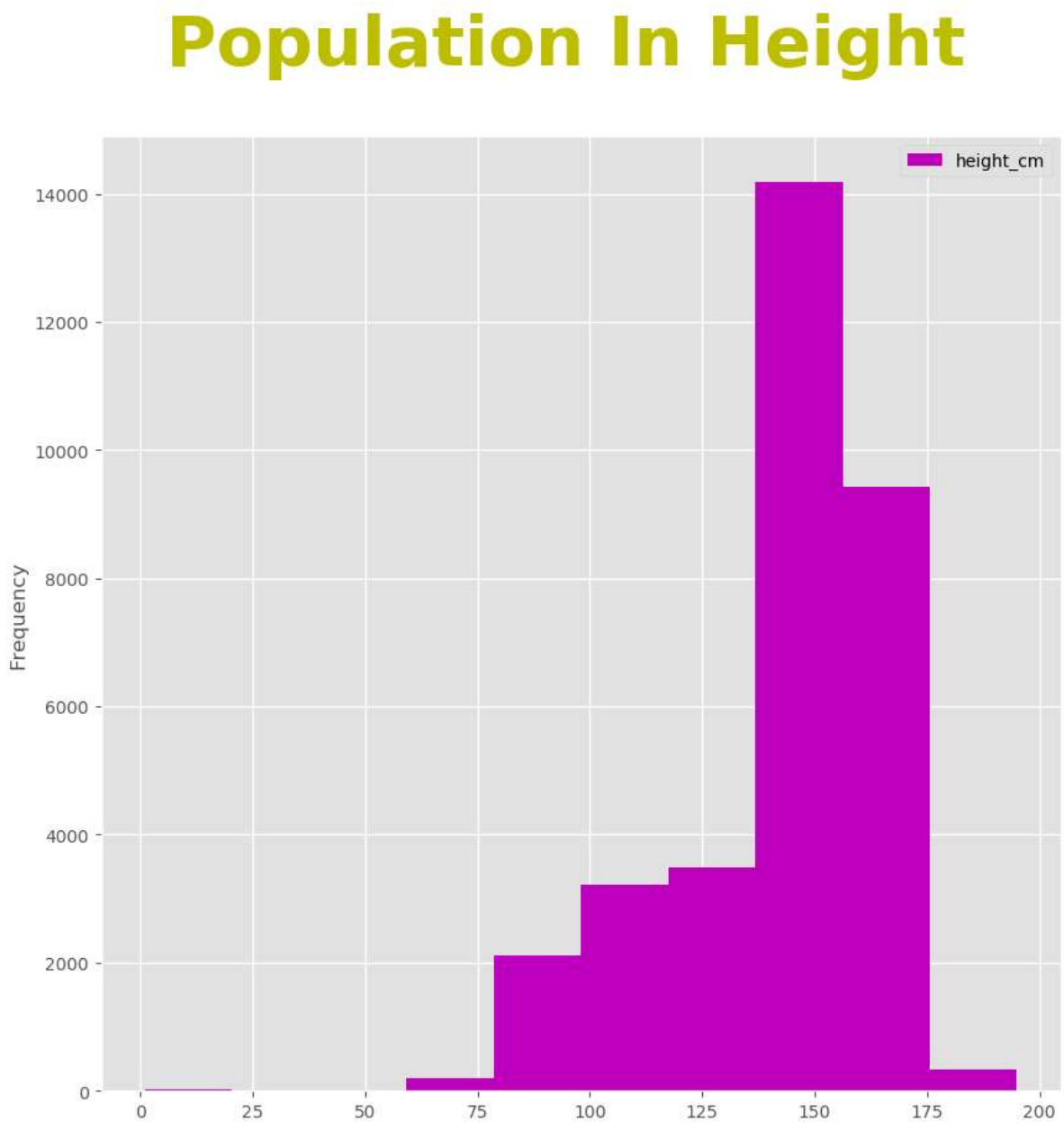
In [24]:
```
# plotting the personid and age
```

In [25]:
```
bihar.plot(x='personid',
           y='age',
           kind="scatter",
           color='tomato',
           figsize=(10,10),
           label='Id');
plt.title=('Population')
plt.legend()
plt.style.use('_classic_test_patch')
plt.suptitle('Population In Age',
             fontsize=40,
             fontweight='bold',
             color='m');
```



In [26]:
```
# plotting the data set in person id and height
```

In [27]:
```
bihar.plot(x='personid',y='height_cm',
           kind="hist",
          figsize=(10,10),
          color='m')
plt.suptitle('Population In Height',
             fontsize=40,
             fontweight='bold',
             color='y');
```

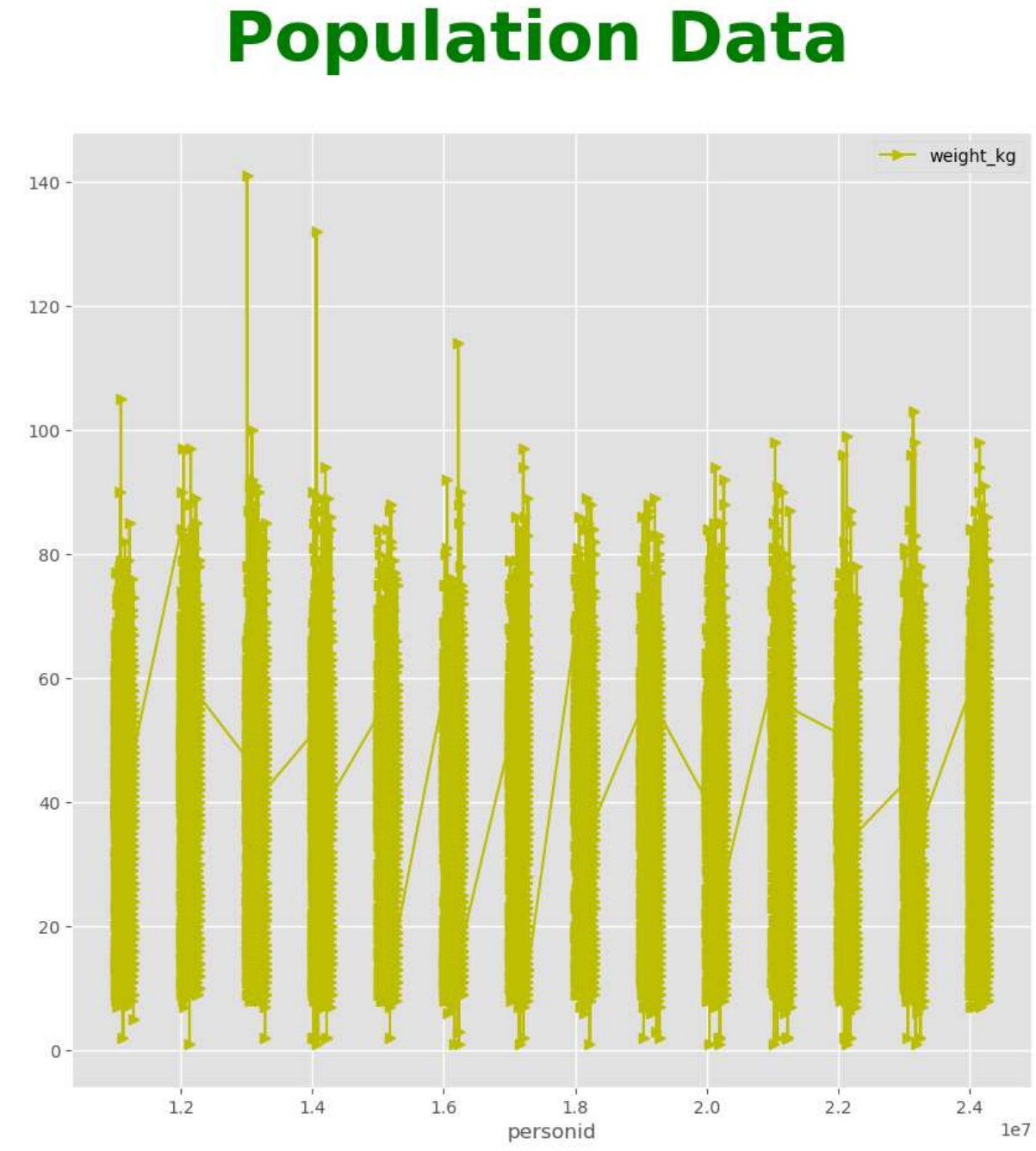


In [28]:
```
plt.style.available
```

Out[28]:
```
['Solarize_Light2',
 '_classic_test_patch',
 '_mpl-gallery',
 '_mpl-gallery-nogrid',
 'bmh',
 'classic',
 'dark_background',
 'fast',
 'fivethirtyeight',
 'ggplot',
 'grayscale',
 'seaborn-v0_8',
 'seaborn-v0_8-bright',
 'seaborn-v0_8-colorblind',
 'seaborn-v0_8-dark',
 'seaborn-v0_8-dark-palette',
 'seaborn-v0_8-darkgrid',
 'seaborn-v0_8-deep',
 'seaborn-v0_8-muted',
 'seaborn-v0_8-notebook',
 'seaborn-v0_8-paper',
 'seaborn-v0_8-pastel',
 'seaborn-v0_8-poster',
 'seaborn-v0_8-talk',
 'seaborn-v0_8-ticks',
 'seaborn-v0_8-white',
 'seaborn-v0_8-whitegrid',
 'tableau-colorblind10']
```

In [29]:
```
# plotting the id and weight
```

In [30]:
```
bihar.plot(x='personid',
           y='weight_kg',
           figsize=(10,10),
           color='y',
           marker='>');
plt.suptitle('Population Data',
              fontsize='40',
              fontweight='bold',
             color='g');
```



In [31]:
```
# plottintg the pyplot of the age of Data
```

pyplot of the data set

In [32]:
```
# import the data set
```

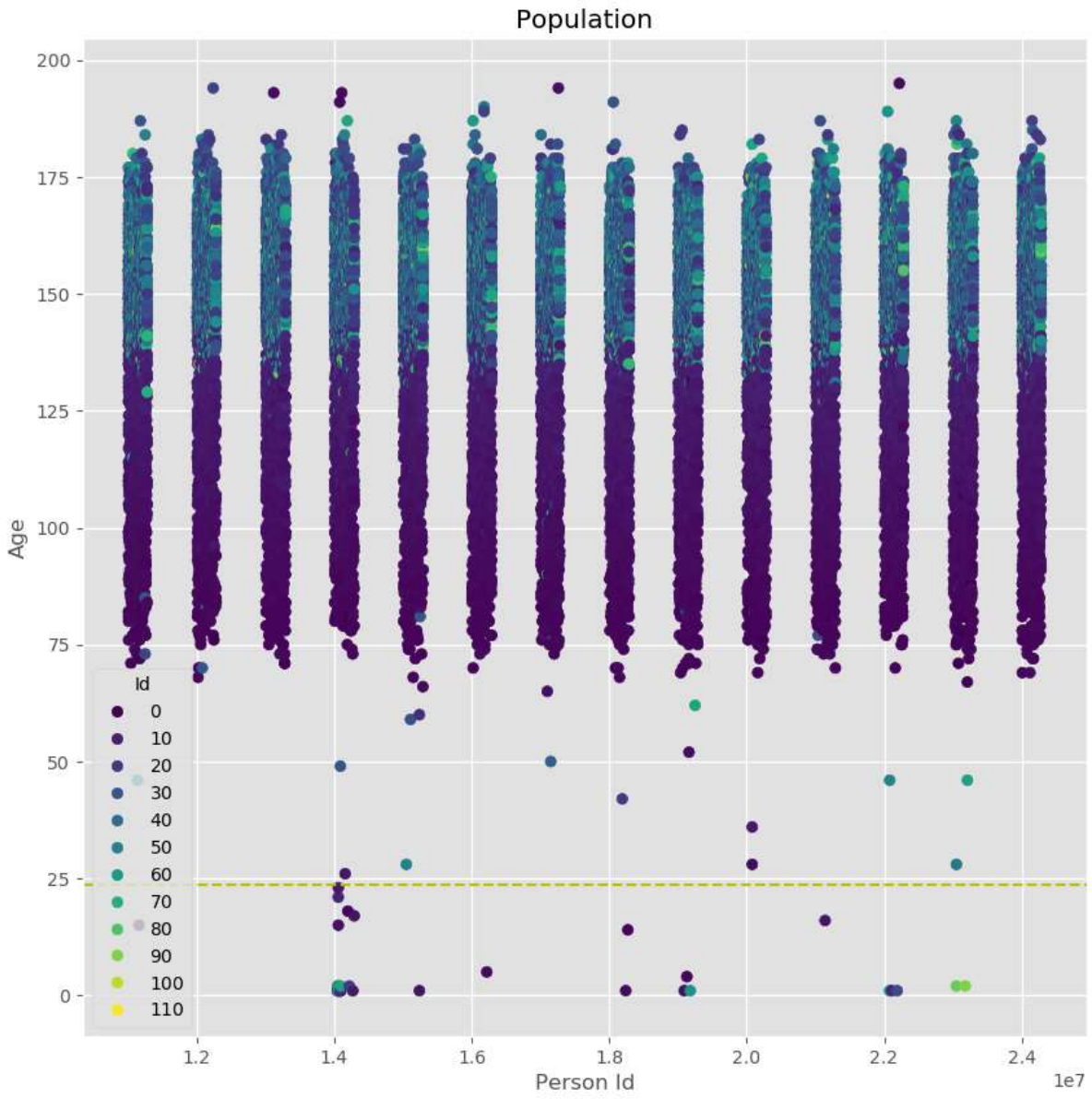
In [33]:
```
bihar
```

Out[33]:

	personid	female	adult	age	height_cm	weight_kg
0	11010101	0	1	70	164	48
1	11010102	0	1	32	157	44
2	11010103	1	1	28	150	37
3	11010104	0	0	12	146	30
4	11010105	1	0	11	135	30
...	...	...	...	...	...	...
39548	24282403	0	1	38	156	60
39549	24282404	1	1	28	148	46
39550	24282405	1	0	7	116	17
39551	24282406	0	0	6	101	13
39552	24282407	1	0	3	91	12

33001 rows × 6 columns

In [34]:
```
fig,figure=plt.subplots(figsize=(10,10))
# adding the data set in plot
axis=plt.scatter(x=bihar['personid'],
                 y=bihar['height_cm'],
                 c=bihar['age'])
# set the data set
figure.set(title='Population',
           xlabel='Person Id',
           ylabel='Age')
# add the legend data
figure.legend(*axis.legend_elements(),title='Id')
#add the mean data
figure.axhline( y=bihar['height_cm'].std(),
                linestyle='--',
               color='y')
# plot style
plt.style.use('fivethirtyeight')
#suptitle
fig.suptitle('Population Data in set',
             fontsize='20',
            color='y')
fig.savefig('Population Data png')
axis;
```



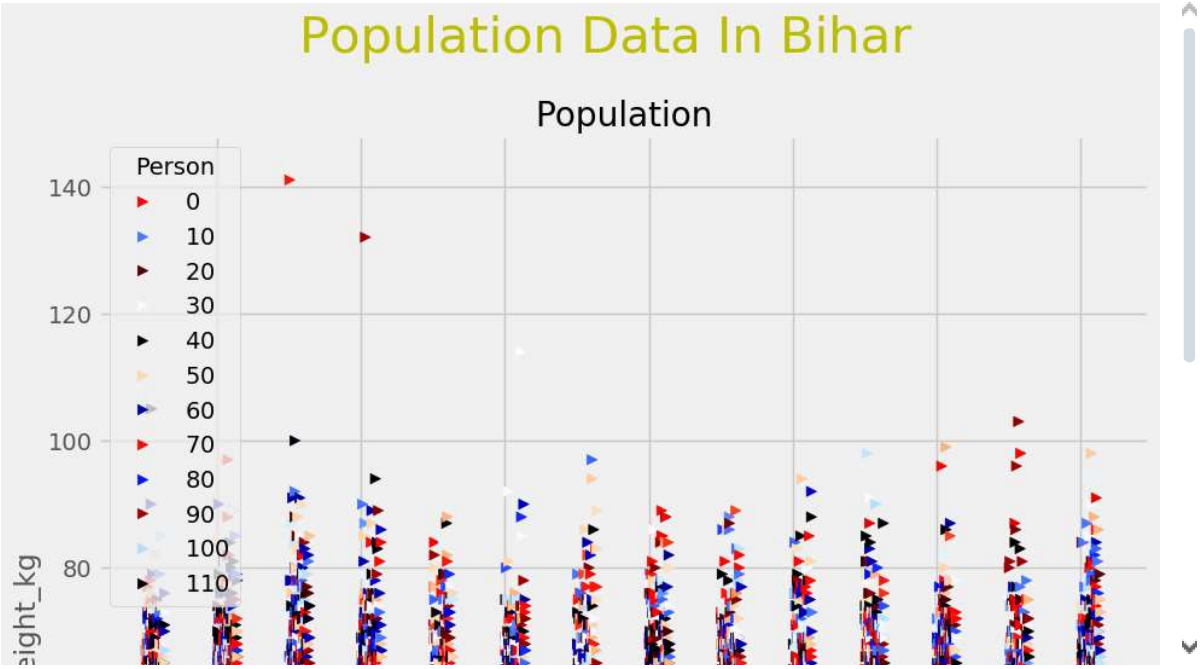
In [35]:
```
# plotting the weight
```

In [36]:
```
bihar.head(2)
```

Out[36]:

	personid	female	adult	age	height_cm	weight_kg
0	11010101	0	1	70	164	48
1	11010102	0	1	32	157	44

In [37]:
```
fig,figure=plt.subplots(figsize=(10,10))
axis=plt.scatter(x=bihar['personid'],
                 y=bihar['weight_kg'],
                 c=bihar['age'],
                 cmap='flag',
                 marker='>')
# set the data
figure.set(title='Population',
           xlabel='Person Id',
           ylabel='Weight_kg')
figure.legend(*axis.legend_elements(),title='Person')
# add the mean value
figure.axhline(y=bihar['weight_kg'].mean(),linestyle='--',color='c',linewidth=
#plotting the data
plt.style.use('fivethirtyeight')
fig.suptitle('Population Data In Bihar',
             fontsize='30',
             fontweight='20',
             color='y')
fig.savefig('population Data in Bihar ')
axis;
```



In [38]:
```
plt.style.available
```

Out[38]:
```
['Solarize_Light2',
 '_classic_test_patch',
 '_mpl-gallery',
 '_mpl-gallery-nogrid',
 'bmh',
 'classic',
 'dark_background',
 'fast',
 'fivethirtyeight',
 'ggplot',
 'grayscale',
 'seaborn-v0_8',
 'seaborn-v0_8-bright',
 'seaborn-v0_8-colorblind',
 'seaborn-v0_8-dark',
 'seaborn-v0_8-dark-palette',
 'seaborn-v0_8-darkgrid',
 'seaborn-v0_8-deep',
 'seaborn-v0_8-muted',
 'seaborn-v0_8-notebook',
 'seaborn-v0_8-paper',
 'seaborn-v0_8-pastel',
 'seaborn-v0_8-poster',
 'seaborn-v0_8-talk',
 'seaborn-v0_8-ticks',
 'seaborn-v0_8-white',
 'seaborn-v0_8-whitegrid',
 'tableau-colorblind10']
```

In [ ]:

In [ ]:

In [ ]: