```
import pandas as pd

df = pd.read_csv('diabetes (1).csv')

print(df.to_string())
```

```
     140            5    106.0          82.0           30      0.0  39.5                    0.286  38.0        0
     141            2    108.0          52.0           26     63.0  32.5                    0.318  22.0        0
     142           10    108.0          66.0            0      0.0  32.4                    0.272  42.0        1
     143            4    154.0          62.0           31    284.0  32.8                    0.237  23.0        0
     144            0    102.0          75.0           23      0.0   0.0                    0.572  21.0        0
     145            9     57.0          80.0           37      0.0  32.8                    0.096  41.0        0
     146            2    106.0          64.0           35    119.0  30.5                    1.400  34.0        0
     147            5    147.0          78.0            0      0.0  33.7                    0.218  65.0        0
     148            2     90.0          70.0           17      0.0  27.3                    0.085  22.0        0
     149           10    129.0          76.0           28    122.0  35.9                    0.280  39.0        0
     150            7    133.0          88.0           15    155.0  32.4                    0.262  37.0        0
     151            7    161.0          86.0            0      0.0  30.4                    0.165  47.0        1
     152            2    108.0          80.0            0      0.0  27.0                    0.259  52.0        1
     153            7    136.0          74.0           26    135.0  26.0                    0.647  51.0        0
     154            5    155.0          84.0           44    545.0  38.7                    0.619  34.0        0
     155            1    119.0          86.0           39    220.0  45.6                    0.808  29.0        1
     156            4     96.0          56.0           17     49.0  20.8                    0.340  26.0        0
     157            5    108.0          72.0           43     75.0  36.1                    0.263  33.0        0
     158            0     78.0          88.0           29     40.0  36.9                    0.434  21.0        0
     159            2    106.0          56.0           27    165.0  29.0                    0.426  22.0        0
     160            2    105.0          75.0            0      0.0  23.3                    0.560   NaN        0
     161            4     95.0          60.0           32      0.0  35.4                    0.284  28.0        0
     162            0    126.0          86.0           27    120.0  27.4                    0.515  21.0        0
     163            8     65.0          72.0           23      0.0  32.0                    0.600  42.0        0
     164            2     99.0           NaN           17    160.0  36.6                    0.453  21.0        0
     165            1    102.0          74.0            0      0.0  39.5                    0.293  42.0        1
     166           11    120.0          80.0           37    150.0  42.3                    0.785  48.0        1
     167            3    102.0          44.0           20     94.0  30.8                    0.400  26.0        0
     168            1    109.0          58.0           18    116.0  28.5                    0.219  22.0        0
     169            9    140.0          94.0            0      0.0  32.7                    0.734  45.0        1
     170           13    153.0          88.0           37    140.0  40.6                    1.174  39.0        0
     171           12    100.0          84.0           33    105.0  30.0                    0.488   NaN        0
     172            1    147.0          94.0           41      0.0  49.3                    0.358  27.0        1
     173            1     81.0          74.0           41     57.0  46.3                    1.096  32.0        0
     174            3    187.0           NaN           22    200.0  36.4                    0.408  36.0        1
     175            6    162.0          62.0            0      0.0  24.3                    0.178  50.0        1
     176            4    136.0          70.0            0      0.0  31.2                    1.182  22.0        1
     177            1    121.0          78.0           39     74.0  39.0                    0.261   NaN        0
     178            3    108.0          62.0           24      0.0  26.0                    0.223  25.0        0
     179            0    181.0           NaN           44    510.0  43.3                    0.222  26.0        1
     180            8    154.0          78.0           32      0.0  32.4                    0.443  45.0        1
     181            1    128.0          88.0           39    110.0  36.5                    1.057  37.0        1
     182            7    137.0          90.0           41      0.0  32.0                    0.391   NaN        0
     183            0    123.0          72.0            0      0.0  36.3                    0.258  52.0        1
     184            1    106.0          76.0            0      0.0  37.5                    0.197  26.0        0
     185            6    190.0           NaN            0      0.0  35.5                    0.278  66.0        1
     186            2     88.0          58.0           26     16.0  28.4                    0.766  22.0        0
     187            9    170.0          74.0           31      0.0  44.0                    0.403  43.0        1
     188            9     89.0          62.0            0      0.0  22.5                    0.142   NaN        0
     189           10    101.0          76.0           48    180.0  32.9                    0.171  63.0        0
     190            2    122.0           NaN           27      0.0  36.8                    0.340  27.0        0
     191            5    121.0          72.0           23    112.0  26.2                    0.245  30.0        0
     192            1    126.0          60.0            0      0.0   NaN                    0.349  47.0        1
     193            1     93.0          70.0           31      0.0  30.4                    0.315   NaN        0
     194            7    137.0          90.0           41      0.0  32.0                    0.391   NaN        0
     195            1    121.0          78.0           39     74.0  39.0                    0.261   NaN        0
     196            0    123.0          72.0            0      0.0  36.3                    0.258  52.0        1
     197            9    170.0          74.0           31      0.0  44.0                    0.403  43.0        1
```

```
df.dropna(inplace = True)

print(df.to_string())
```

145	9	57.0	80.0	37	0.0	32.8	0.096	41.0	0
146	2	106.0	64.0	35	119.0	30.5	1.400	34.0	0
147	5	147.0	78.0	0	0.0	33.7	0.218	65.0	0
148	2	90.0	70.0	17	0.0	27.3	0.085	22.0	0
149	10	129.0	76.0	28	122.0	35.9	0.280	39.0	0
150	7	133.0	88.0	15	155.0	32.4	0.262	37.0	0
151	7	161.0	86.0	0	0.0	30.4	0.165	47.0	1
152	2	108.0	80.0	0	0.0	27.0	0.259	52.0	1
153	7	136.0	74.0	26	135.0	26.0	0.647	51.0	0
154	5	155.0	84.0	44	545.0	38.7	0.619	34.0	0
155	1	119.0	86.0	39	220.0	45.6	0.808	29.0	1
156	4	96.0	56.0	17	49.0	20.8	0.340	26.0	0
157	5	108.0	72.0	43	75.0	36.1	0.263	33.0	0
158	0	78.0	88.0	29	40.0	36.9	0.434	21.0	0
159	2	106.0	56.0	27	165.0	29.0	0.426	22.0	0
161	4	95.0	60.0	32	0.0	35.4	0.284	28.0	0
162	0	126.0	86.0	27	120.0	27.4	0.515	21.0	0
163	8	65.0	72.0	23	0.0	32.0	0.600	42.0	0
165	1	102.0	74.0	0	0.0	39.5	0.293	42.0	1
166	11	120.0	80.0	37	150.0	42.3	0.785	48.0	1
167	3	102.0	44.0	20	94.0	30.8	0.400	26.0	0
168	1	109.0	58.0	18	116.0	28.5	0.219	22.0	0
169	9	140.0	94.0	0	0.0	32.7	0.734	45.0	1
170	13	153.0	88.0	37	140.0	40.6	1.174	39.0	0
172	1	147.0	94.0	41	0.0	49.3	0.358	27.0	1
173	1	81.0	74.0	41	57.0	46.3	1.096	32.0	0
175	6	162.0	62.0	0	0.0	24.3	0.178	50.0	1
176	4	136.0	70.0	0	0.0	31.2	1.182	22.0	1
178	3	108.0	62.0	24	0.0	26.0	0.223	25.0	0
180	8	154.0	78.0	32	0.0	32.4	0.443	45.0	1
181	1	128.0	88.0	39	110.0	36.5	1.057	37.0	1
183	0	123.0	72.0	0	0.0	36.3	0.258	52.0	1
184	1	106.0	76.0	0	0.0	37.5	0.197	26.0	0
186	2	88.0	58.0	26	16.0	28.4	0.766	22.0	0
187	9	170.0	74.0	31	0.0	44.0	0.403	43.0	1
189	10	101.0	76.0	48	180.0	32.9	0.171	63.0	0
191	5	121.0	72.0	23	112.0	26.2	0.245	30.0	0
196	0	123.0	72.0	0	0.0	36.3	0.258	52.0	1
197	9	170.0	74.0	31	0.0	44.0	0.403	43.0	1

Python provides numerous libraries for data analysis and visualization mainly numpy, pandas, matplotlib, seaborn etc.

```
#Read first five element
df.head()
```

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Ou
0	6	148.0	72.0	35	0.0	33.6	0.627	50.0	
1	1	85.0	66.0	29	0.0	26.6	0.351	31.0	
2	8	183.0	64.0	0	0.0	23.3	0.672	32.0	
3	1	89.0	66.0	23	94.0	28.1	0.167	21.0	
4	0	137.0	40.0	35	168.0	43.1	2.288	33.0	

```
#Countthe number of rows and columns
df.shape
```

```
(167, 9)
```

```
#To computes various summary statistics
df.describe()
```

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFu
count	167.000000	167.000000	167.000000	167.000000	167.000000	167.000000	167.
mean	4.377246	118.574850	69.844311	20.161677	71.958084	31.883234	0.
std	3.485973	31.166171	18.401472	15.892833	113.646427	8.930054	0.
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.
25%	1.000000	99.500000	62.000000	0.000000	0.000000	27.400000	0.
50%	4.000000	115.000000	72.000000	23.000000	0.000000	32.600000	0.
75%	7.000000	140.500000	80.000000	32.000000	117.500000	36.600000	0.
max	15.000000	196.000000	122.000000	60.000000	846.000000	55.000000	2.

```
# computes numerical data ranks
df.rank()
```

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age
0	114.5	139.0	83.5	138.0	44.0	97.5	131.0	144.0
1	32.0	22.5	58.0	109.0	44.0	36.5	82.0	80.5
2	139.5	164.0	48.5	27.0	44.0	19.0	136.0	85.5
3	32.0	27.5	58.0	81.5	112.0	47.0	13.0	6.0
4	8.5	120.5	8.0	138.0	145.5	152.0	167.0	91.0
...	...	...	...	...	...	...	...	...
187	147.5	157.0	97.5	120.5	44.0	156.5	95.0	130.5
189	155.5	46.5	111.5	163.0	150.0	90.0	14.0	165.0
191	102.5	93.0	83.5	81.5	123.0	34.0	38.5	75.5
196	8.5	98.0	83.5	27.0	44.0	122.5	47.5	150.0
197	147.5	157.0	97.5	120.5	44.0	156.5	95.0	130.5

167 rows × 9 columns

Plotting

Pandas uses the plot() method to create diagrams.

We can use Pyplot, a submodule of the Matplotlib library to visualize the diagram on the screen.

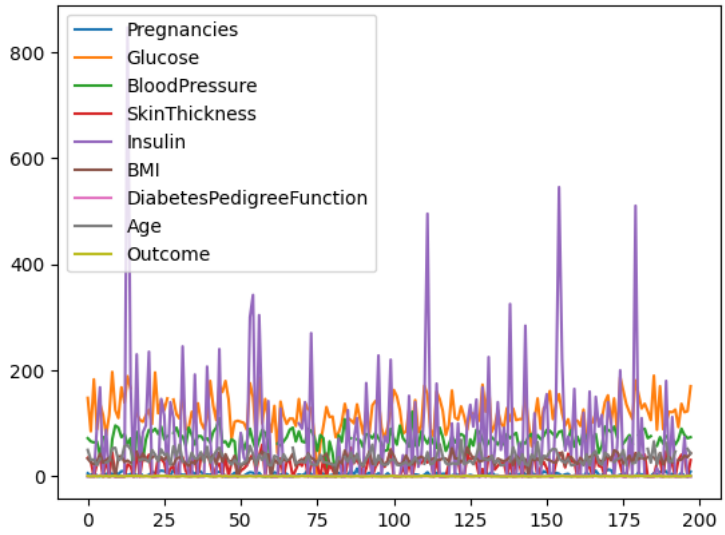
Import pyplot from Matplotlib and visualize our DataFrame:

```
import pandas as pd
import matplotlib.pyplot as plt

df = pd.read_csv('diabetes (1).csv')

df.plot()

plt.show()
```



Scatter Plot

Specify that you want a scatter plot with the kind argument:

kind = 'scatter'

A scatter plot needs an x- and a y-axis.

In the example use "Age" for the x-axis and "BloodPressure" for the y-axis.

Include the x and y arguments like this:

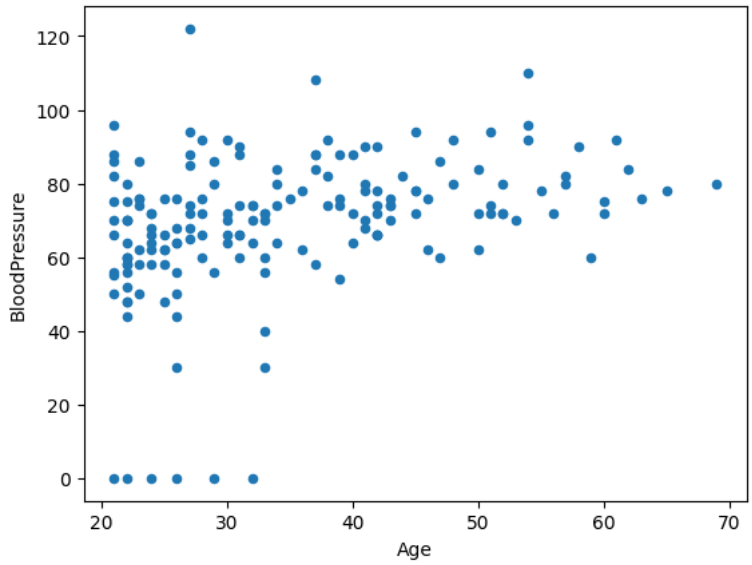
x = 'Age', y = 'BloodPressure'

```
import pandas as pd
import matplotlib.pyplot as plt

df = pd.read_csv('diabetes (1).csv')
```

```
df.plot(kind = 'scatter', x = 'Age', y = 'BloodPressure')

plt.show()
```

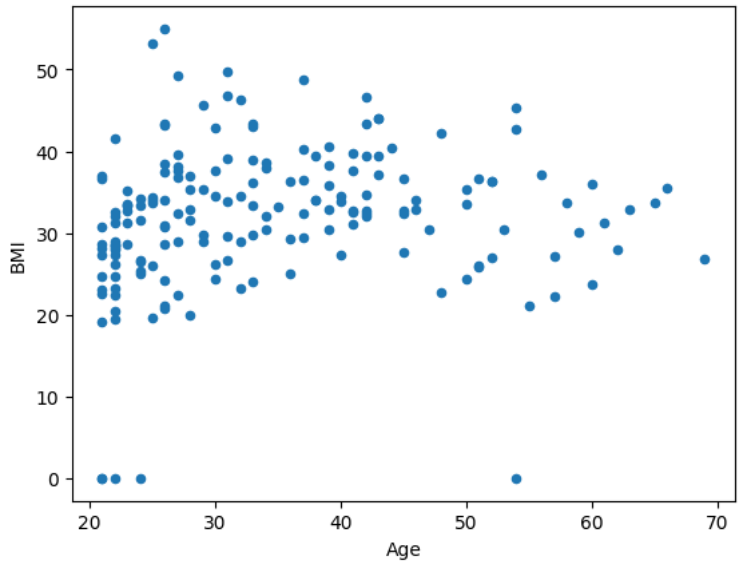



```
import pandas as pd
import matplotlib.pyplot as plt

df = pd.read_csv('diabetes (1).csv')

df.plot(kind = 'scatter', x = 'Age', y = 'BMI')

plt.show()
```



Histogram

Use the kind argument to specify that you want a histogram:

kind = 'hist'

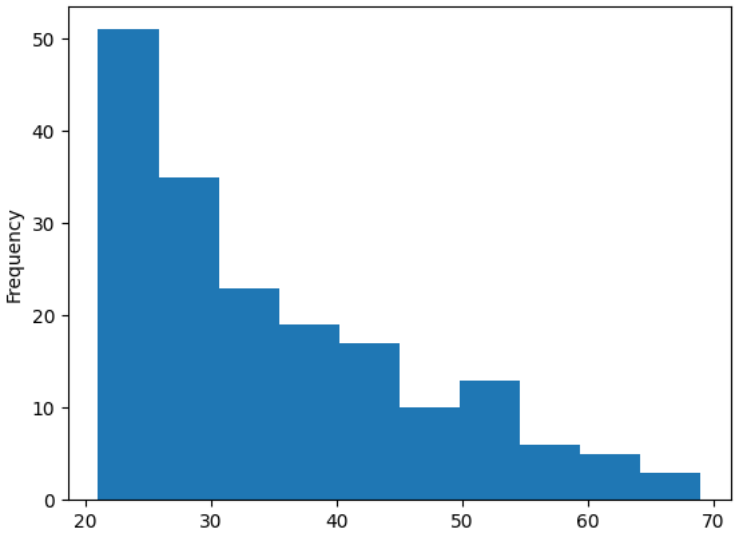
A histogram needs only one column.

```
import pandas as pd
import matplotlib.pyplot as plt

df = pd.read_csv('diabetes (1).csv')

df["Age"].plot(kind = 'hist')
```

```
<Axes: ylabel='Frequency'>
```



✓ 0s completed at 10:24 AM