6m3ttbflm

April 29, 2025

[18]:
```
import numpy as np
import pandas as pd
```

[19]:
```
# Read the dataset
data = pd.read_csv('Insurance.csv')
```

[20]:
```
data.head()
```

[20]:

	age	sex	bmi	children	smoker	region	charges
0	19	female	27.900	0	yes	southwest	16884.92400
1	18	male	33.770	1	no	southeast	1725.55230
2	28	male	33.000	3	no	southeast	4449.46200
3	33	male	22.705	0	no	northwest	21984.47061
4	32	male	28.880	0	no	northwest	3866.85520

[21]:
```
data.tail()
```

[21]:

	age	sex	bmi	children	smoker	region	charges
34	28	male	36.400	1	yes	southwest	51194.55914
35	19	male	20.425	0	no	northwest	1625.43375
36	62	female	32.965	3	no	northwest	15612.19335
37	26	male	20.800	0	no	southwest	2302.30000
38	35	male	36.670	1	yes	northeast	39774.27630

[22]:
```
# Calculate summary statistics grouped by a categorical variable
categorical_variable = 'age'  # Replace with your actual column name
numeric_variable = 'charges'  # Replace with your actual column name

summary_stats = data.groupby(categorical_variable)[numeric_variable].describe()

print(summary_stats)
```

```
     count          mean           std          min           25%  \
age
18     3.0   1686.917717    531.643312   1137.01100   1431.281650
19     4.0   6258.847937   7220.361576   1625.43375   1784.286188
22     1.0  35585.576000           NaN  35585.57600  35585.576000
23     3.0   2332.402233    477.280335   1826.84300   2111.007275
```

25	1.0	2721.320800	NaN	2721.32080	2721.320800
26	1.0	2302.300000	NaN	2302.30000	2302.300000
27	1.0	39611.757700	NaN	39611.75770	39611.757700
28	2.0	27822.010570	33053.775175	4449.46200	16135.736285
30	2.0	20493.601500	23113.716252	4149.73600	12321.668750
31	2.0	21233.810800	24716.477999	3756.62160	12495.216200
32	1.0	3866.855200	NaN	3866.85520	3866.855200
33	1.0	21984.470610	NaN	21984.47061	21984.470610
34	1.0	37701.876800	NaN	37701.87680	37701.876800
35	1.0	39774.276300	NaN	39774.27630	39774.276300
37	3.0	6630.606017	572.717356	6203.90175	6305.156225
46	1.0	8240.589600	NaN	8240.58960	8240.589600
52	1.0	10797.336200	NaN	10797.33620	10797.336200
55	1.0	12268.632250	NaN	12268.63225	12268.632250
56	2.0	10846.551400	345.303434	10602.38500	10724.468200
59	1.0	14001.133800	NaN	14001.13380	14001.133800
60	2.0	21075.991935	11097.538864	13228.84695	17152.419442
62	2.0	21710.459225	8624.250307	15612.19335	18661.326288
63	2.0	14110.966525	482.061032	13770.09790	13940.532213

age	50%	75%	max
18	1725.552300	1961.871075	2198.18985
19	3262.517000	7737.078750	16884.92400
22	35585.576000	35585.576000	35585.57600
23	2395.171550	2585.181850	2775.19215
25	2721.320800	2721.320800	2721.32080
26	2302.300000	2302.300000	2302.30000
27	39611.757700	39611.757700	39611.75770
28	27822.010570	39508.284855	51194.55914
30	20493.601500	28665.534250	36837.46700
31	21233.810800	29972.405400	38711.00000
32	3866.855200	3866.855200	3866.85520
33	21984.470610	21984.470610	21984.47061
34	37701.876800	37701.876800	37701.87680
35	39774.276300	39774.276300	39774.27630
37	6406.410700	6843.958150	7281.50560
46	8240.589600	8240.589600	8240.58960
52	10797.336200	10797.336200	10797.33620
55	12268.632250	12268.632250	12268.63225
56	10846.551400	10968.634600	11090.71780
59	14001.133800	14001.133800	14001.13380
60	21075.991935	24999.564428	28923.13692
62	21710.459225	24759.592163	27808.72510
63	14110.966525	14281.400838	14451.83515

[23]:
```
# Extract the desired statistics for each category
desired_statistics = ['mean', 'median', 'min', 'max', 'std']

# Create a list of numeric values for each response to the categorical variable
category_values = summary_stats.index.tolist()

# Print the summary statistics
print(summary_stats)
print("Category values:", category_values)
```

```
     count          mean           std          min           25%  \
age
18     3.0   1686.917717    531.643312   1137.01100   1431.281650
19     4.0   6258.847937   7220.361576   1625.43375   1784.286188
22     1.0  35585.576000           NaN  35585.57600  35585.576000
23     3.0   2332.402233    477.280335   1826.84300   2111.007275
25     1.0   2721.320800           NaN   2721.32080   2721.320800
26     1.0   2302.300000           NaN   2302.30000   2302.300000
27     1.0  39611.757700           NaN  39611.75770  39611.757700
28     2.0  27822.010570  33053.775175   4449.46200  16135.736285
30     2.0  20493.601500  23113.716252   4149.73600  12321.668750
31     2.0  21233.810800  24716.477999   3756.62160  12495.216200
32     1.0   3866.855200           NaN   3866.85520   3866.855200
33     1.0  21984.470610           NaN  21984.47061  21984.470610
34     1.0  37701.876800           NaN  37701.87680  37701.876800
35     1.0  39774.276300           NaN  39774.27630  39774.276300
37     3.0   6630.606017    572.717356   6203.90175   6305.156225
46     1.0   8240.589600           NaN   8240.58960   8240.589600
52     1.0  10797.336200           NaN  10797.33620  10797.336200
55     1.0  12268.632250           NaN  12268.63225  12268.632250
56     2.0  10846.551400    345.303434  10602.38500  10724.468200
59     1.0  14001.133800           NaN  14001.13380  14001.133800
60     2.0  21075.991935  11097.538864  13228.84695  17152.419442
62     2.0  21710.459225   8624.250307  15612.19335  18661.326288
63     2.0  14110.966525    482.061032  13770.09790  13940.532213

              50%           75%          max
age
18    1725.552300   1961.871075   2198.18985
19    3262.517000   7737.078750  16884.92400
22   35585.576000  35585.576000  35585.57600
23    2395.171550   2585.181850   2775.19215
25    2721.320800   2721.320800   2721.32080
26    2302.300000   2302.300000   2302.30000
27   39611.757700  39611.757700  39611.75770
28   27822.010570  39508.284855  51194.55914
30   20493.601500  28665.534250  36837.46700
```

```
31   21233.810800  29972.405400  38711.00000
32    3866.855200   3866.855200   3866.85520
33   21984.470610  21984.470610  21984.47061
34   37701.876800  37701.876800  37701.87680
35   39774.276300  39774.276300  39774.27630
37    6406.410700   6843.958150   7281.50560
46    8240.589600   8240.589600   8240.58960
52   10797.336200  10797.336200  10797.33620
55   12268.632250  12268.632250  12268.63225
56   10846.551400  10968.634600  11090.71780
59   14001.133800  14001.133800  14001.13380
60   21075.991935  24999.564428  28923.13692
62   21710.459225  24759.592163  27808.72510
63   14110.966525  14281.400838  14451.83515
Category values: [18, 19, 22, 23, 25, 26, 27, 28, 30, 31, 32, 33, 34, 35, 37,
46, 52, 55, 56, 59, 60, 62, 63]
```

Part 2

[24]:
```
df = pd.read_csv('Iris.csv')
```

[25]:
```
df.head()
```

[25]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.1	3.5	1.4	0.2	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa
3	4	4.6	3.1	1.5	0.2	Iris-setosa
4	5	5.0	3.6	1.4	0.2	Iris-setosa

[26]:
```
df.tail()
```

[26]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	\
145	146	6.7	3.0	5.2	2.3	
146	147	6.3	2.5	5.0	1.9	
147	148	6.5	3.0	5.2	2.0	
148	149	6.2	3.4	5.4	2.3	
149	150	5.9	3.0	5.1	1.8	

	Species
145	Iris-virginica
146	Iris-virginica
147	Iris-virginica
148	Iris-virginica
149	Iris-virginica

[27]:
```
# Filter dataset for 'Iris-setosa', 'Iris-versicolor', and 'Iris-virginica'
setosa_df = df[df['Species'] == 'Iris-setosa']
```

```
versicolor_df = df[df['Species'] == 'Iris-versicolor']
virginica_df = df[df['Species'] == 'Iris-virginica']
```

[28]:
```
print(setosa_df)
```

```
    Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm      Species
0    1            5.1           3.5            1.4           0.2  Iris-setosa
1    2            4.9           3.0            1.4           0.2  Iris-setosa
2    3            4.7           3.2            1.3           0.2  Iris-setosa
3    4            4.6           3.1            1.5           0.2  Iris-setosa
4    5            5.0           3.6            1.4           0.2  Iris-setosa
5    6            5.4           3.9            1.7           0.4  Iris-setosa
6    7            4.6           3.4            1.4           0.3  Iris-setosa
7    8            5.0           3.4            1.5           0.2  Iris-setosa
8    9            4.4           2.9            1.4           0.2  Iris-setosa
9   10            4.9           3.1            1.5           0.1  Iris-setosa
10  11            5.4           3.7            1.5           0.2  Iris-setosa
11  12            4.8           3.4            1.6           0.2  Iris-setosa
12  13            4.8           3.0            1.4           0.1  Iris-setosa
13  14            4.3           3.0            1.1           0.1  Iris-setosa
14  15            5.8           4.0            1.2           0.2  Iris-setosa
15  16            5.7           4.4            1.5           0.4  Iris-setosa
16  17            5.4           3.9            1.3           0.4  Iris-setosa
17  18            5.1           3.5            1.4           0.3  Iris-setosa
18  19            5.7           3.8            1.7           0.3  Iris-setosa
19  20            5.1           3.8            1.5           0.3  Iris-setosa
20  21            5.4           3.4            1.7           0.2  Iris-setosa
21  22            5.1           3.7            1.5           0.4  Iris-setosa
22  23            4.6           3.6            1.0           0.2  Iris-setosa
23  24            5.1           3.3            1.7           0.5  Iris-setosa
24  25            4.8           3.4            1.9           0.2  Iris-setosa
25  26            5.0           3.0            1.6           0.2  Iris-setosa
26  27            5.0           3.4            1.6           0.4  Iris-setosa
27  28            5.2           3.5            1.5           0.2  Iris-setosa
28  29            5.2           3.4            1.4           0.2  Iris-setosa
29  30            4.7           3.2            1.6           0.2  Iris-setosa
30  31            4.8           3.1            1.6           0.2  Iris-setosa
31  32            5.4           3.4            1.5           0.4  Iris-setosa
32  33            5.2           4.1            1.5           0.1  Iris-setosa
33  34            5.5           4.2            1.4           0.2  Iris-setosa
34  35            4.9           3.1            1.5           0.1  Iris-setosa
35  36            5.0           3.2            1.2           0.2  Iris-setosa
36  37            5.5           3.5            1.3           0.2  Iris-setosa
37  38            4.9           3.1            1.5           0.1  Iris-setosa
38  39            4.4           3.0            1.3           0.2  Iris-setosa
39  40            5.1           3.4            1.5           0.2  Iris-setosa
40  41            5.0           3.5            1.3           0.3  Iris-setosa
```

```
41  42            4.5           2.3            1.3           0.3  Iris-setosa
42  43            4.4           3.2            1.3           0.2  Iris-setosa
43  44            5.0           3.5            1.6           0.6  Iris-setosa
44  45            5.1           3.8            1.9           0.4  Iris-setosa
45  46            4.8           3.0            1.4           0.3  Iris-setosa
46  47            5.1           3.8            1.6           0.2  Iris-setosa
47  48            4.6           3.2            1.4           0.2  Iris-setosa
48  49            5.3           3.7            1.5           0.2  Iris-setosa
49  50            5.0           3.3            1.4           0.2  Iris-setosa
```

[29]:
```
print(versicolor_df)
```

```
     Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm  \
50   51            7.0           3.2            4.7           1.4
51   52            6.4           3.2            4.5           1.5
52   53            6.9           3.1            4.9           1.5
53   54            5.5           2.3            4.0           1.3
54   55            6.5           2.8            4.6           1.5
55   56            5.7           2.8            4.5           1.3
56   57            6.3           3.3            4.7           1.6
57   58            4.9           2.4            3.3           1.0
58   59            6.6           2.9            4.6           1.3
59   60            5.2           2.7            3.9           1.4
60   61            5.0           2.0            3.5           1.0
61   62            5.9           3.0            4.2           1.5
62   63            6.0           2.2            4.0           1.0
63   64            6.1           2.9            4.7           1.4
64   65            5.6           2.9            3.6           1.3
65   66            6.7           3.1            4.4           1.4
66   67            5.6           3.0            4.5           1.5
67   68            5.8           2.7            4.1           1.0
68   69            6.2           2.2            4.5           1.5
69   70            5.6           2.5            3.9           1.1
70   71            5.9           3.2            4.8           1.8
71   72            6.1           2.8            4.0           1.3
72   73            6.3           2.5            4.9           1.5
73   74            6.1           2.8            4.7           1.2
74   75            6.4           2.9            4.3           1.3
75   76            6.6           3.0            4.4           1.4
76   77            6.8           2.8            4.8           1.4
77   78            6.7           3.0            5.0           1.7
78   79            6.0           2.9            4.5           1.5
79   80            5.7           2.6            3.5           1.0
80   81            5.5           2.4            3.8           1.1
81   82            5.5           2.4            3.7           1.0
82   83            5.8           2.7            3.9           1.2
83   84            6.0           2.7            5.1           1.6
84   85            5.4           3.0            4.5           1.5
```

85	86	6.0	3.4	4.5	1.6
86	87	6.7	3.1	4.7	1.5
87	88	6.3	2.3	4.4	1.3
88	89	5.6	3.0	4.1	1.3
89	90	5.5	2.5	4.0	1.3
90	91	5.5	2.6	4.4	1.2
91	92	6.1	3.0	4.6	1.4
92	93	5.8	2.6	4.0	1.2
93	94	5.0	2.3	3.3	1.0
94	95	5.6	2.7	4.2	1.3
95	96	5.7	3.0	4.2	1.2
96	97	5.7	2.9	4.2	1.3
97	98	6.2	2.9	4.3	1.3
98	99	5.1	2.5	3.0	1.1
99	100	5.7	2.8	4.1	1.3

	Species
50	Iris-versicolor
51	Iris-versicolor
52	Iris-versicolor
53	Iris-versicolor
54	Iris-versicolor
55	Iris-versicolor
56	Iris-versicolor
57	Iris-versicolor
58	Iris-versicolor
59	Iris-versicolor
60	Iris-versicolor
61	Iris-versicolor
62	Iris-versicolor
63	Iris-versicolor
64	Iris-versicolor
65	Iris-versicolor
66	Iris-versicolor
67	Iris-versicolor
68	Iris-versicolor
69	Iris-versicolor
70	Iris-versicolor
71	Iris-versicolor
72	Iris-versicolor
73	Iris-versicolor
74	Iris-versicolor
75	Iris-versicolor
76	Iris-versicolor
77	Iris-versicolor
78	Iris-versicolor
79	Iris-versicolor
80	Iris-versicolor

```
81  Iris-versicolor
82  Iris-versicolor
83  Iris-versicolor
84  Iris-versicolor
85  Iris-versicolor
86  Iris-versicolor
87  Iris-versicolor
88  Iris-versicolor
89  Iris-versicolor
90  Iris-versicolor
91  Iris-versicolor
92  Iris-versicolor
93  Iris-versicolor
94  Iris-versicolor
95  Iris-versicolor
96  Iris-versicolor
97  Iris-versicolor
98  Iris-versicolor
99  Iris-versicolor
```

[30]:
```
print(virginica_df)
```

```
      Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm  \
100  101            6.3           3.3            6.0           2.5
101  102            5.8           2.7            5.1           1.9
102  103            7.1           3.0            5.9           2.1
103  104            6.3           2.9            5.6           1.8
104  105            6.5           3.0            5.8           2.2
105  106            7.6           3.0            6.6           2.1
106  107            4.9           2.5            4.5           1.7
107  108            7.3           2.9            6.3           1.8
108  109            6.7           2.5            5.8           1.8
109  110            7.2           3.6            6.1           2.5
110  111            6.5           3.2            5.1           2.0
111  112            6.4           2.7            5.3           1.9
112  113            6.8           3.0            5.5           2.1
113  114            5.7           2.5            5.0           2.0
114  115            5.8           2.8            5.1           2.4
115  116            6.4           3.2            5.3           2.3
116  117            6.5           3.0            5.5           1.8
117  118            7.7           3.8            6.7           2.2
118  119            7.7           2.6            6.9           2.3
119  120            6.0           2.2            5.0           1.5
120  121            6.9           3.2            5.7           2.3
121  122            5.6           2.8            4.9           2.0
122  123            7.7           2.8            6.7           2.0
123  124            6.3           2.7            4.9           1.8
124  125            6.7           3.3            5.7           2.1
```

125	126	7.2	3.2	6.0	1.8
126	127	6.2	2.8	4.8	1.8
127	128	6.1	3.0	4.9	1.8
128	129	6.4	2.8	5.6	2.1
129	130	7.2	3.0	5.8	1.6
130	131	7.4	2.8	6.1	1.9
131	132	7.9	3.8	6.4	2.0
132	133	6.4	2.8	5.6	2.2
133	134	6.3	2.8	5.1	1.5
134	135	6.1	2.6	5.6	1.4
135	136	7.7	3.0	6.1	2.3
136	137	6.3	3.4	5.6	2.4
137	138	6.4	3.1	5.5	1.8
138	139	6.0	3.0	4.8	1.8
139	140	6.9	3.1	5.4	2.1
140	141	6.7	3.1	5.6	2.4
141	142	6.9	3.1	5.1	2.3
142	143	5.8	2.7	5.1	1.9
143	144	6.8	3.2	5.9	2.3
144	145	6.7	3.3	5.7	2.5
145	146	6.7	3.0	5.2	2.3
146	147	6.3	2.5	5.0	1.9
147	148	6.5	3.0	5.2	2.0
148	149	6.2	3.4	5.4	2.3
149	150	5.9	3.0	5.1	1.8

	Species
100	Iris-virginica
101	Iris-virginica
102	Iris-virginica
103	Iris-virginica
104	Iris-virginica
105	Iris-virginica
106	Iris-virginica
107	Iris-virginica
108	Iris-virginica
109	Iris-virginica
110	Iris-virginica
111	Iris-virginica
112	Iris-virginica
113	Iris-virginica
114	Iris-virginica
115	Iris-virginica
116	Iris-virginica
117	Iris-virginica
118	Iris-virginica
119	Iris-virginica
120	Iris-virginica

```
121  Iris-virginica
122  Iris-virginica
123  Iris-virginica
124  Iris-virginica
125  Iris-virginica
126  Iris-virginica
127  Iris-virginica
128  Iris-virginica
129  Iris-virginica
130  Iris-virginica
131  Iris-virginica
132  Iris-virginica
133  Iris-virginica
134  Iris-virginica
135  Iris-virginica
136  Iris-virginica
137  Iris-virginica
138  Iris-virginica
139  Iris-virginica
140  Iris-virginica
141  Iris-virginica
142  Iris-virginica
143  Iris-virginica
144  Iris-virginica
145  Iris-virginica
146  Iris-virginica
147  Iris-virginica
148  Iris-virginica
149  Iris-virginica
```

[31]:
```
# Calculate statistical details for each species
setosa_stats = setosa_df.describe()
versicolor_stats = versicolor_df.describe()
virginica_stats = virginica_df.describe()
```

[32]:
```
print("Statistical details for 'Iris-setosa':")
print(setosa_stats)
```

```
Statistical details for 'Iris-setosa':
             Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm
count  50.00000       50.00000     50.000000      50.000000      50.00000
mean   25.50000        5.00600      3.418000       1.464000       0.24400
std    14.57738        0.35249      0.381024       0.173511       0.10721
min     1.00000        4.30000      2.300000       1.000000       0.10000
25%    13.25000        4.80000      3.125000       1.400000       0.20000
50%    25.50000        5.00000      3.400000       1.500000       0.20000
75%    37.75000        5.20000      3.675000       1.575000       0.30000
max    50.00000        5.80000      4.400000       1.900000       0.60000
```

[33]:
```
print("\nStatistical details for 'Iris-versicolor':")
print(versicolor_stats)
```

```
Statistical details for 'Iris-versicolor':
              Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm
count   50.00000      50.000000     50.000000      50.000000     50.000000
mean    75.50000       5.936000      2.770000       4.260000      1.326000
std     14.57738       0.516171      0.313798       0.469911      0.197753
min     51.00000       4.900000      2.000000       3.000000      1.000000
25%     63.25000       5.600000      2.525000       4.000000      1.200000
50%     75.50000       5.900000      2.800000       4.350000      1.300000
75%     87.75000       6.300000      3.000000       4.600000      1.500000
max    100.00000       7.000000      3.400000       5.100000      1.800000
```

[34]:
```
print("\nStatistical details for 'Iris-virginica':")
print(virginica_stats)
```

```
Statistical details for 'Iris-virginica':
              Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm
count   50.00000       50.00000     50.000000      50.000000      50.00000
mean   125.50000        6.58800      2.974000       5.552000       2.02600
std     14.57738        0.63588      0.322497       0.551895       0.27465
min    101.00000        4.90000      2.200000       4.500000       1.40000
25%    113.25000        6.22500      2.800000       5.100000       1.80000
50%    125.50000        6.50000      3.000000       5.550000       2.00000
75%    137.75000        6.90000      3.175000       5.875000       2.30000
max    150.00000        7.90000      3.800000       6.900000       2.50000
```

[ ]:

[ ]:

[ ]:

[ ]:

[ ]: