

```
import numpy as np
import pandas as pd
```

```
from google.colab import files
uploaded = files.upload()
```

[Choose Files](#) imports-85.data.txt

- **imports-85.data.txt**(text/plain) - 25936 bytes, last modified: 3/15/2023 - 100% done
Saving imports-85.data.txt to imports-85.data.txt

```
df = pd.read_csv('imports-85.data.txt')
```

df

| | 3 | ? | alfa-romero | gas | std | two | convertible | rwd | front | 88.60 | ... | 130 | mpfi | 3.47 | 2.68 | 9.00 | 111 | 5000 | 21 | 27 | 1 |
|-----|-----|-----|-------------|--------|-------|------|-------------|-----|-------|-------|-----|-----|------|------|------|------|-----|------|-----|-----|-----|
| 0 | 3 | ? | alfa-romero | gas | std | two | convertible | rwd | front | 88.6 | ... | 130 | mpfi | 3.47 | 2.68 | 9.0 | 111 | 5000 | 21 | 27 | 1 |
| 1 | 1 | ? | alfa-romero | gas | std | two | hatchback | rwd | front | 94.5 | ... | 152 | mpfi | 2.68 | 3.47 | 9.0 | 154 | 5000 | 19 | 26 | 1 |
| 2 | 2 | 164 | audi | gas | std | four | sedan | fwd | front | 99.8 | ... | 109 | mpfi | 3.19 | 3.40 | 10.0 | 102 | 5500 | 24 | 30 | 1 |
| 3 | 2 | 164 | audi | gas | std | four | sedan | 4wd | front | 99.4 | ... | 136 | mpfi | 3.19 | 3.40 | 8.0 | 115 | 5500 | 18 | 22 | 1 |
| 4 | 2 | ? | audi | gas | std | two | sedan | fwd | front | 99.8 | ... | 136 | mpfi | 3.19 | 3.40 | 8.5 | 110 | 5500 | 19 | 25 | 1 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 199 | -1 | 95 | volvo | gas | std | four | sedan | rwd | front | 109.1 | ... | 141 | mpfi | 3.78 | 3.15 | 9.5 | 114 | 5400 | 23 | 28 | 1 |
| 200 | -1 | 95 | volvo | gas | turbo | four | sedan | rwd | front | 109.1 | ... | 141 | mpfi | 3.78 | 3.15 | 8.7 | 160 | 5300 | 19 | 25 | 1 |
| 201 | -1 | 95 | volvo | gas | std | four | sedan | rwd | front | 109.1 | ... | 173 | mpfi | 3.58 | 2.87 | 8.8 | 134 | 5500 | 18 | 23 | 2 |
| 202 | -1 | 95 | volvo | diesel | turbo | four | sedan | rwd | front | 109.1 | ... | 145 | idi | 3.01 | 3.40 | 23.0 | 106 | 4800 | 26 | 27 | 2 |
| 203 | -1 | 95 | volvo | gas | turbo | four | sedan | rwd | front | 109.1 | ... | 141 | mpfi | 3.78 | 3.15 | 9.5 | 114 | 5400 | 19 | 25 | 2 |

df.head()

| | 3 | ? | alfa-romero | gas | std | two | convertible | rwd | front | 88.60 | ... | 130 | mpfi | 3.47 | 2.68 | 9.00 | 111 | 5000 | 21 | 27 | 13495 |
|---|---|-----|-------------|-----|-----|------|-------------|-----|-------|-------|-----|-----|------|------|------|------|-----|------|----|----|-------|
| 0 | 3 | ? | alfa-romero | gas | std | two | convertible | rwd | front | 88.6 | ... | 130 | mpfi | 3.47 | 2.68 | 9.0 | 111 | 5000 | 21 | 27 | 16500 |
| 1 | 1 | ? | alfa-romero | gas | std | two | hatchback | rwd | front | 94.5 | ... | 152 | mpfi | 2.68 | 3.47 | 9.0 | 154 | 5000 | 19 | 26 | 16500 |
| 2 | 2 | 164 | audi | gas | std | four | sedan | fwd | front | 99.8 | ... | 109 | mpfi | 3.19 | 3.40 | 10.0 | 102 | 5500 | 24 | 30 | 13950 |
| 3 | 2 | 164 | audi | gas | std | four | sedan | 4wd | front | 99.4 | ... | 136 | mpfi | 3.19 | 3.40 | 8.0 | 115 | 5500 | 18 | 22 | 17450 |
| 4 | 2 | ? | audi | gas | std | two | sedan | fwd | front | 99.8 | | 136 | mpfi | 3.19 | 3.40 | 8.5 | 110 | 5500 | 19 | 25 | 15250 |

df.tail()

| | 3 | ? | alfa-romero | gas | std | two | convertible | rwd | front | 88.60 | ... | 130 | mpfi | 3.47 | 2.68 | 9.00 | 111 | 5000 | 21 | 27 | 13 |
|-----|----|----|-------------|--------|-------|------|-------------|-----|-------|-------|-----|-----|------|------|------|------|-----|------|----|----|----|
| 199 | -1 | 95 | volvo | gas | std | four | sedan | rwd | front | 109.1 | ... | 141 | mpfi | 3.78 | 3.15 | 9.5 | 114 | 5400 | 23 | 28 | 16 |
| 200 | -1 | 95 | volvo | gas | turbo | four | sedan | rwd | front | 109.1 | ... | 141 | mpfi | 3.78 | 3.15 | 8.7 | 160 | 5300 | 19 | 25 | 19 |
| 201 | -1 | 95 | volvo | gas | std | four | sedan | rwd | front | 109.1 | ... | 173 | mpfi | 3.58 | 2.87 | 8.8 | 134 | 5500 | 18 | 23 | 21 |
| 202 | -1 | 95 | volvo | diesel | turbo | four | sedan | rwd | front | 109.1 | ... | 145 | idi | 3.01 | 3.40 | 23.0 | 106 | 4800 | 26 | 27 | 22 |
| 203 | -1 | 95 | volvo | gas | turbo | four | sedan | rwd | front | 109.1 | ... | 141 | mpfi | 3.78 | 3.15 | 9.5 | 114 | 5400 | 19 | 25 | 22 |

5 rows × 26 columns

df.isnull()

| | 3 | ? | alfa-romero | gas | std | two | convertible | rwd | front | 88.60 | ... | 130 | mpfi | 3.47 | 2.68 | 9.00 | 111 | 500 |
|-----|-------|-------|-------------|-------|-------|-------|-------------|-------|-------|-------|-----|-------|-------|-------|-------|-------|-------|-------|
| 0 | False | False | False | False | False | False | False | False | False | False | ... | False | False | False | False | False | False | False |
| 1 | False | False | False | False | False | False | False | False | False | False | ... | False | False | False | False | False | False | False |
| 2 | False | False | False | False | False | False | False | False | False | False | ... | False | False | False | False | False | False | False |
| 3 | False | False | False | False | False | False | False | False | False | False | ... | False | False | False | False | False | False | False |
| 4 | False | False | False | False | False | False | False | False | False | False | ... | False | False | False | False | False | False | False |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 199 | False | False | False | False | False | False | False | False | False | False | ... | False | False | False | False | False | False | False |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |

```
df.describe()
```

| | 3 | 88.60 | 168.80 | 64.10 | 48.80 | 2548 | 130 | 9.00 | 21 | 27 |
|--------------|------------|------------|------------|------------|------------|-------------|------------|------------|------------|------------|
| count | 204.000000 | 204.000000 | 204.000000 | 204.000000 | 204.000000 | 204.000000 | 204.000000 | 204.000000 | 204.000000 | 204.000000 |
| mean | 0.823529 | 98.806373 | 174.075000 | 65.916667 | 53.749020 | 2555.602941 | 126.892157 | 10.148137 | 25.240196 | 30.769608 |
| std | 1.239035 | 5.994144 | 12.362123 | 2.146716 | 2.424901 | 521.960820 | 41.744569 | 3.981000 | 6.551513 | 6.898337 |
| min | -2.000000 | 86.600000 | 141.100000 | 60.300000 | 47.800000 | 1488.000000 | 61.000000 | 7.000000 | 13.000000 | 16.000000 |
| 25% | 0.000000 | 94.500000 | 166.300000 | 64.075000 | 52.000000 | 2145.000000 | 97.000000 | 8.575000 | 19.000000 | 25.000000 |
| 50% | 1.000000 | 97.000000 | 173.200000 | 65.500000 | 54.100000 | 2414.000000 | 119.500000 | 9.000000 | 24.000000 | 30.000000 |
| 75% | 2.000000 | 102.400000 | 183.200000 | 66.900000 | 55.500000 | 2939.250000 | 142.000000 | 9.400000 | 30.000000 | 34.500000 |
| max | 3.000000 | 120.900000 | 208.100000 | 72.300000 | 59.800000 | 4066.000000 | 326.000000 | 23.000000 | 49.000000 | 54.000000 |

```
df.dropna()
```

| | | | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-------------|-----|-----|------|-------------|-----|-------|-------|-----|-----|------|------|------|------|-----|------|-----|-----|-----|
| | 3 | ? | alfa-romero | gas | std | two | convertible | rwd | front | 88.60 | ... | 130 | mpfi | 3.47 | 2.68 | 9.00 | 111 | 5000 | 21 | 27 | 1 |
| 0 | 3 | ? | alfa-romero | gas | std | two | convertible | rwd | front | 88.6 | ... | 130 | mpfi | 3.47 | 2.68 | 9.0 | 111 | 5000 | 21 | 27 | 1 |
| 1 | 1 | ? | alfa-romero | gas | std | two | hatchback | rwd | front | 94.5 | ... | 152 | mpfi | 2.68 | 3.47 | 9.0 | 154 | 5000 | 19 | 26 | 1 |
| 2 | 2 | 164 | audi | gas | std | four | sedan | fwd | front | 99.8 | ... | 109 | mpfi | 3.19 | 3.40 | 10.0 | 102 | 5500 | 24 | 30 | 1 |
| 3 | 2 | 164 | audi | gas | std | four | sedan | 4wd | front | 99.4 | ... | 136 | mpfi | 3.19 | 3.40 | 8.0 | 115 | 5500 | 18 | 22 | 1 |
| 4 | 2 | ? | audi | gas | std | two | sedan | fwd | front | 99.8 | ... | 136 | mpfi | 3.19 | 3.40 | 8.5 | 110 | 5500 | 19 | 25 | 1 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 199 | -1 | 95 | volvo | gas | std | four | sedan | rwd | front | 109.1 | ... | 141 | mpfi | 3.78 | 3.15 | 9.5 | 114 | 5400 | 23 | 28 | 1 |

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 204 entries, 0 to 203
Data columns (total 26 columns):
#   Column          Non-Null Count  Dtype
---  -
0    3                204 non-null    int64
1    ?                204 non-null    object
2    alfa-romero      204 non-null    object
3    gas              204 non-null    object
4    std              204 non-null    object
5    two              204 non-null    object
6    convertible      204 non-null    object
7    rwd              204 non-null    object
8    front            204 non-null    object
9    88.60            204 non-null    float64
10   168.80           204 non-null    float64
11   64.10            204 non-null    float64
12   48.80            204 non-null    float64
13   2548             204 non-null    int64
14   dohc             204 non-null    object
15   four             204 non-null    object
16   130              204 non-null    int64
17   mpfi             204 non-null    object
18   3.47             204 non-null    object
```

```

19  2.68          204 non-null  object
20  9.00          204 non-null  float64
21  111          204 non-null  object
22  5000          204 non-null  object
23  21           204 non-null  int64
24  27           204 non-null  int64
25  13495        204 non-null  object
dtypes: float64(5), int64(5), object(16)
memory usage: 41.6+ KB

```

```
df = df.astype({'gas':'string'})
```

```
df.info()
```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 204 entries, 0 to 203
Data columns (total 26 columns):
#   Column          Non-Null Count  Dtype
---  -
0   3               204 non-null   int64
1   ?              204 non-null   object
2   alfa-romero    204 non-null   object
3   gas            204 non-null   string
4   std            204 non-null   object
5   two            204 non-null   object
6   convertible    204 non-null   object
7   rwd            204 non-null   object
8   front          204 non-null   object
9   88.60          204 non-null   float64
10  168.80          204 non-null   float64
11  64.10           204 non-null   float64
12  48.80           204 non-null   float64
13  2548            204 non-null   int64
14  dohc            204 non-null   object
15  four            204 non-null   object
16  130             204 non-null   int64
17  mpfi            204 non-null   object
18  3.47            204 non-null   object
19  2.68            204 non-null   object
20  9.00            204 non-null   float64
21  111             204 non-null   object
22  5000            204 non-null   object
23  21              204 non-null   int64

```

```
24 27          204 non-null    int64
25 13495       204 non-null    object
dtypes: float64(5), int64(5), object(15), string(1)
memory usage: 41.6+ KB
```

```
pd.get_dummies(df['std'])
```

| | std | turbo |
|-----|-----|-------|
| 0 | 1 | 0 |
| 1 | 1 | 0 |
| 2 | 1 | 0 |
| 3 | 1 | 0 |
| 4 | 1 | 0 |
| ... | ... | ... |
| 199 | 1 | 0 |
| 200 | 0 | 1 |
| 201 | 1 | 0 |
| 202 | 0 | 1 |
| 203 | 0 | 1 |

204 rows × 2 columns

```
df.dtypes
```

```
3          int64
?          object
alfa-romero  object
gas          string
std          object
two          object
convertible  object
rwd          object
front        object
88.60        float64
```

| | |
|--------|---------|
| 168.80 | float64 |
| 64.10 | float64 |
| 48.80 | float64 |
| 2548 | int64 |
| dohc | object |
| four | object |
| 130 | int64 |
| mpfi | object |
| 3.47 | object |
| 2.68 | object |
| 9.00 | float64 |
| 111 | object |
| 5000 | object |
| 21 | int64 |
| 27 | int64 |
| 13495 | object |
| dtype: | object |