In [1]:

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

CSVファイルを開く前にテキストエディタで中身を調べる。
Before reading the csv file, inspect the file using text editor.

In [2]:

```
# note the options
df=pd.read_csv('ai_mid1.csv', skiprows=2, delimiter=';', header=0)
```

In [3]:

```
df.head()
```

Out[3]:

	price	horsepower	width	height
0	13495.0	111.0	64.1	48.8
1	16500.0	111.0	64.1	48.8
2	16500.0	154.0	65.5	52.4
3	13950.0	102.0	66.2	54.3
4	17450.0	115.0	66.4	54.3

In [4]:

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 205 entries, 0 to 204
Data columns (total 4 columns):
 #   Column      Non-Null Count  Dtype  
---  ------      --------------  -----  
 0   price       201 non-null    float64
 1   horsepower  203 non-null    float64
 2   width       205 non-null    float64
 3   height      205 non-null    float64
dtypes: float64(4)
memory usage: 6.5 KB
```

Ans.1

205

widthの平均値
mean value of width

In [5]:

```
df['width'].mean()
```

Out[5]:

65.90780487804878

Ans.2

65.9

欠損値を調べる。
Check NaN contents

In [6]:

```
df.isnull().sum(axis=0)
```

Out[6]:

```
price         4
horsepower    2
width         0
height        0
dtype: int64
```

Ans.3

price, horsepower

Ans.4

In [7]:

```
# use dropna and reset_index
df=df.dropna().reset_index(drop=True)
```

In [8]:

```
df.shape
```

Out[8]:

(199, 4)

馬力/幅を計算して新しく列にする
add a column with "horsepower/width"

In [9]:

```
df['hppw']=df['horsepower']/df['width']
```

最大値を与えるindexを求める。
finding the index whose ratio is the maximum

In [10]:

```
df['hppw'].idxmax()
```

Out[10]:

46

Ans.5

46