1. 数据整理，归纳

```
原始数据说明:
  persons = [
    # firstname,    secondname,  height,   weight,  gender
    [ 'Randy',      'Carter',    '184',    '73',   'male'],
    [ 'Stephanie', 'Smith',     '149',    '52',   'female'],
    ...
  ]
```

```
'''
返回数据
firstnames
{
  'male': ['Randy', 'Jessie', 'David', 'Stephen', 'Jerry',
  'female': ['Stephanie', 'Cynthia', 'Katherine', 'Elizabe
}
heights
{
  'male': ['184', '175', '187', '192', '204', '180', '184'
  'female': ['149', '174', '183', '138', '145', '161', '17
}
'''
```

2. 特征值

例如: heights 在规定区间范围内有多少数值

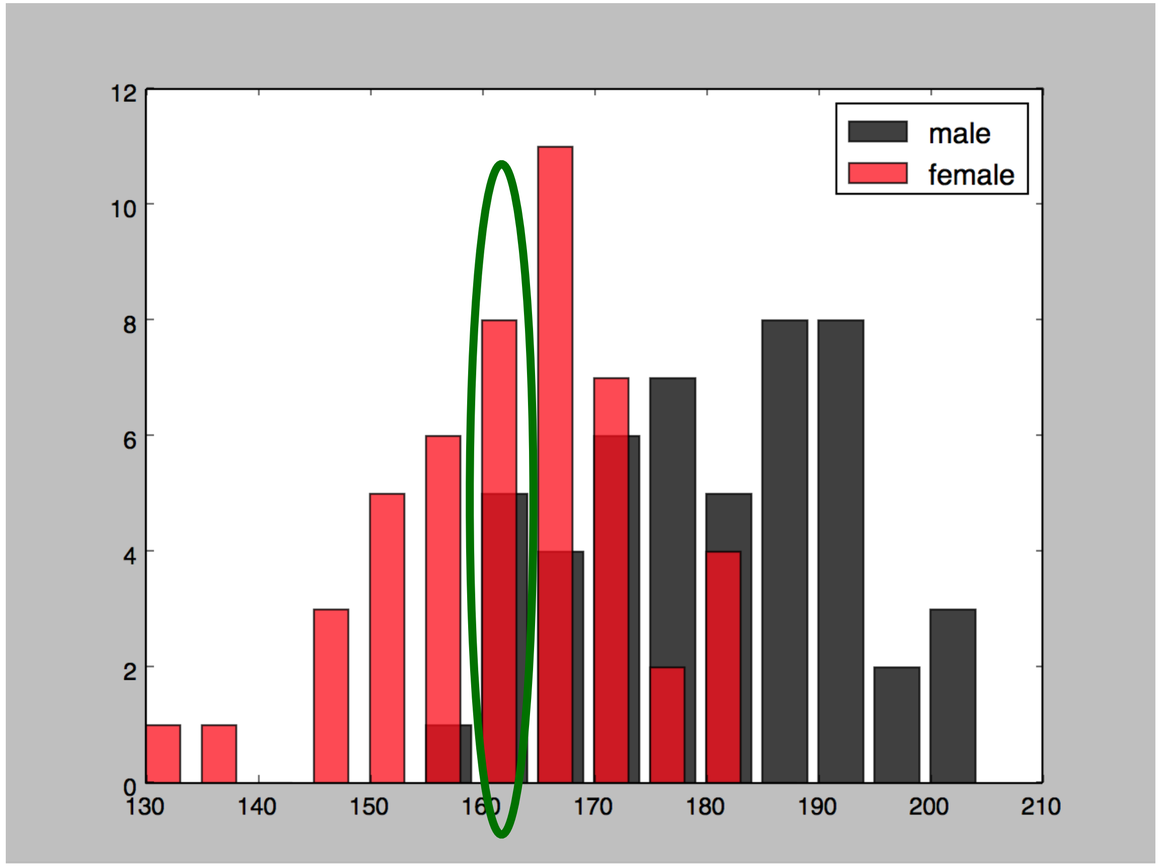
```
'''
原始数据
[ 184 175 187 192 204 180 184 174 177 200 193 189 188 187 187 190 180 155
  201 171 163 191 175 170 178 191 176 168 169 173 180 212 189 174 167 193
  164 171 165 198 185 175 195 164 187 192 175 190 164 161
]

[ 149 174 183 138 145 161 179 162 148 196 163 159 150 170 169 167 168 170
  167 152 155 151 166 165 158 152 159 163 174 131 154 159 159 177 180 164
  163 161 183 170 167 166 180 161 170 172 168 156 167 167
]
'''
```

```
'''
  返回数据
  freq_dict: {160: 5, 195: 2, 165: 4, 200: 3, 170: 6, 175: 7, 180: 5, 185: 8, 155: 1, 190: 8}
  freq_dict: {160: 8, 130: 1, 165: 11, 135: 1, 170: 7, 140: 0, 175: 2, 145: 3, 180: 4, 150: 5,
```

例如: 区间160-165共多少个?

3. 绘图



female: 8
male: 5

```
$ python main.py
('[Feature Data]', 'male', {160: 5, 195: 2, 165: 4, 200: 3, 170: 6, 175: 7, 180: 5, 185: 8, 155: 1, 190: 8}, 49)
('[Feature Data]', 'female', {160: 8, 130: 1, 165: 11, 135: 1, 170: 7, 140: 0, 175: 2, 145: 3, 180: 4, 150: 5, 185: 0, 155: 6}, 48)
```

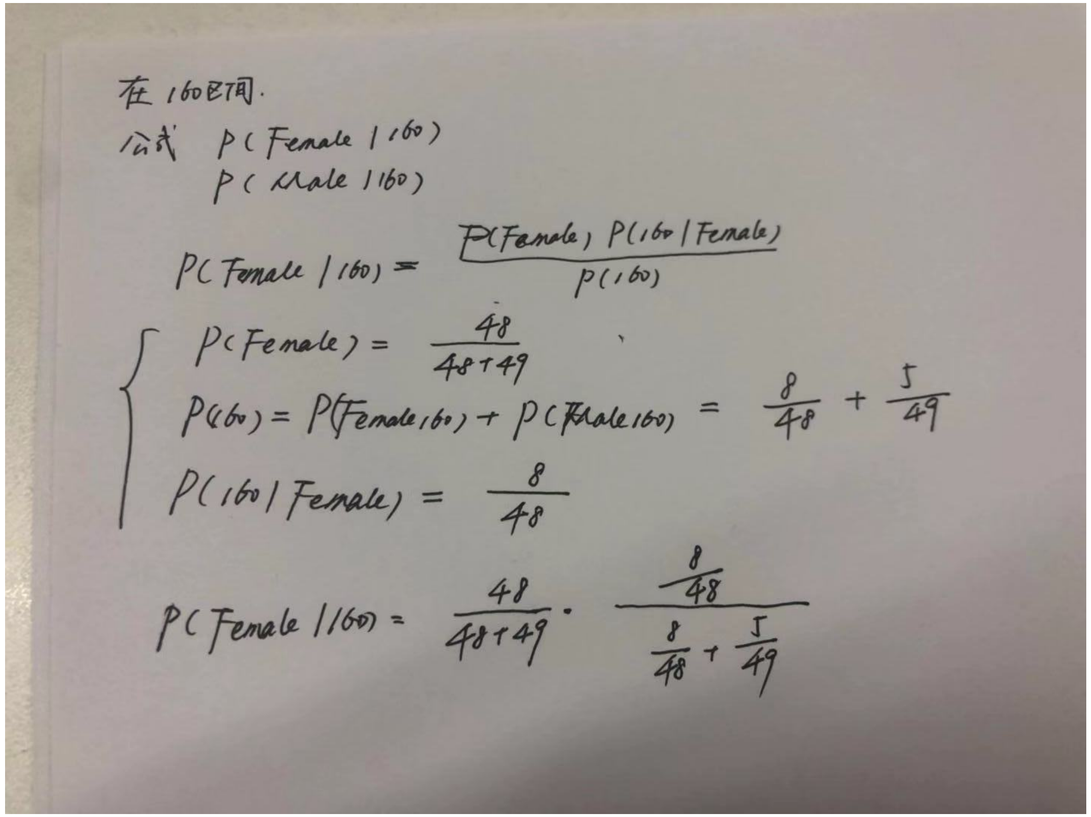
4. 发生概率

限定条件为已知性别, 求在该性别下某高度出现的概率

```
'''
例如在160区间
Female: 8, 总数48
Male: 5, 总数49

P(F160) = 8 / 48 = 0.16666666666666666
P(M160) = 5 / 49 = 0.10204081632653061
P(160) = P(F160) + P(M160) = 0.16666666666666666 + 0.10204081632653061 = 0.26870748299319724

Female 概率: 0.16666666666666666 / 0.26870748299319724 = 0.6202531645569621
Male 概率: 0.10204081632653061 / 0.26870748299319724 = 0.37974683544303806
'''
```



```
'''
返回值
(130, [(0.0, 'male'), (1.0, 'female')])
(135, [(0.0, 'male'), (1.0, 'female')])
(140, [(0.5, 'male'), (0.5, 'female')])
(145, [(0.0, 'male'), (1.0, 'female')])
(150, [(0.0, 'male'), (1.0, 'female')])
(155, [(0.14035087719298245, 'male'), (0.85964912280701755, 'female')])
(160, [(0.37974683544303806, 'male'), (0.62025316455696211, 'female')])
(165, [(0.26265389876880985, 'male'), (0.73734610123119015, 'female')])
(170, [(0.456418383518225, 'male'), (0.54358161648177494, 'female')])
(175, [(0.77419354838709675, 'male'), (0.22580645161290325, 'female')])
(180, [(0.55045871559633031, 'male'), (0.44954128440366975, 'female')])
(185, [(1.0, 'male'), (0.0, 'female')])
(190, [(1.0, 'male'), (0.0, 'female')])
(195, [(1.0, 'male'), (0.0, 'female')])
(200, [(1.0, 'male'), (0.0, 'female')])
(205, [(0.5, 'male'), (0.5, 'female')])
(210, [(0.5, 'male'), (0.5, 'female')])
(215, [(0.5, 'male'), (0.5, 'female')])
'''
```