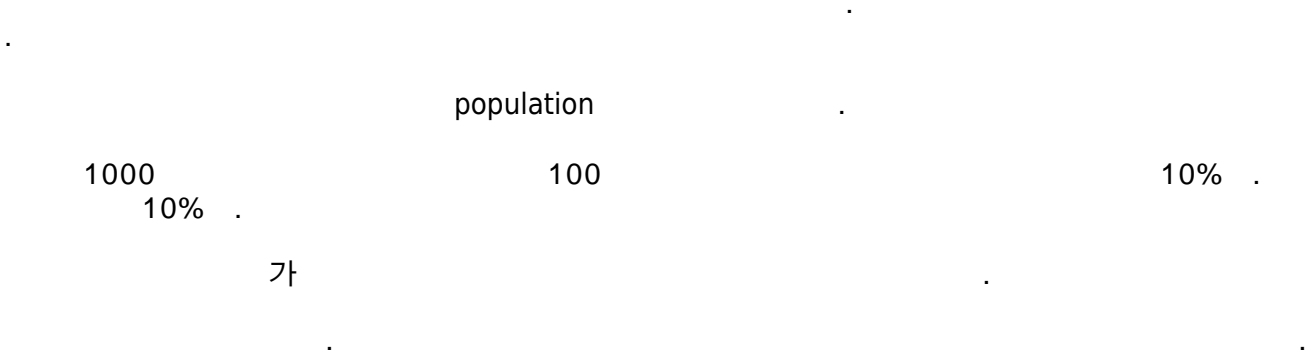


-
- , ,
- z-score
-

?



가 ,

Observational

Experimental

가 가

. 가 , 가 ,

confounding factor .

가 .
'가 ' .

가 가 , 가 '
가' .
.\$\$ s = \frac{\sum(x-\bar{x})^2}{n} \$\$

- s .
- x .
- \bar{x} .
- n .

snippet.python

```
np.std(friends) # == 425.2
```

variation . measure of

z-score

z-score

x z-score . $z = \frac{x - \bar{x}}{s}$.

- x
- \bar{x}
- s

snippet.python

```
z_scores = []

m = np.mean(friends)
s = np.std(friends)

for friend in friends:
    z = (friend - m)/s
    z_scores.append(z)
```

가 가 . -
 . (가 . (가) -
 .) .) - (가)
 correlation coefficients /

snippet.python

```
friends = [109, 1017, 1127, 418, 625, 957, 89, 950, 946, 797, 981, 125,
455, 731, 1640, 485, 1309, 472, 1132, 1773, 906, 531, 742, 621]
happiness = [.8, .6, .3, .6, .6, .4, .8, .5, .4, .3, .3, .6, .2, .8, 1,
.6, .2, .7, .5, .3, .1, 0, .3, 1]

import pandas as pd

df = pd.DataFrame({'friends':friends, 'happiness':happiness})
df.head()
```

```
<div>
<table border="1" class="dataframe">
  <thead>
    <tr style="text-align: right;">
      <th></th>
      <th>friends</th>
      <th>happiness</th>
    </tr>
  </thead>
```

```
<tbody>
  <tr>
    <th>0</th>
    <td>109</td>
    <td>0.8</td>
  </tr>
  <tr>
    <th>1</th>
    <td>1017</td>
    <td>0.6</td>
  </tr>
  <tr>
    <th>2</th>
    <td>1127</td>
    <td>0.3</td>
  </tr>
  <tr>
    <th>3</th>
    <td>418</td>
    <td>0.6</td>
  </tr>
  <tr>
    <th>4</th>
    <td>625</td>
    <td>0.6</td>
  </tr>
</tbody>
</table>
</div>
```

[snippet.python](#)

```
#
df.corr()
```

```
<div>
<table border="1" class="dataframe">
  <thead>
    <tr style="text-align: right;">
      <th></th>
      <th>friends</th>
      <th>happiness</th>
    </tr>
  </thead>
  <tbody>
    <tr>
      <th>friends</th>
      <td>1.000000</td>
      <td>-0.216199</td>
```

```

</tr>
<tr>
  <th>happiness</th>
  <td>-0.216199</td>
  <td>1.000000</td>
</tr>
</tbody>
</table>
</div>

```

snippet.python

```

from sklearn import preprocessing
df_scaled = pd.DataFrame(preprocessing.scale(df),
columns=['friends_scaled', 'happiness_scaled'])
df_scaled.head()

```

```

<div>
<table border="1" class="dataframe">
  <thead>
    <tr style="text-align: right;">
      <th></th>
      <th>friends_scaled</th>
      <th>happiness_scaled</th>
    </tr>
  </thead>
  <tbody>
    <tr>
      <th>0</th>
      <td>-1.599495</td>
      <td>1.153223</td>
    </tr>
    <tr>
      <th>1</th>
      <td>0.536040</td>
      <td>0.394939</td>
    </tr>
    <tr>
      <th>2</th>
      <td>0.794750</td>
      <td>-0.742486</td>
    </tr>
    <tr>
      <th>3</th>
      <td>-0.872755</td>
      <td>0.394939</td>
    </tr>
    <tr>
      <th>4</th>

```

```
<td>-0.385909</td>
<td>0.394939</td>
</tr>
</tbody>
</table>
</div>
```

[snippet.python](#)

```
# 1
within_1_std = df_scaled[(df_scaled['friends_scaled'] <= 1) &
(df_scaled['friends_scaled'] >= -1)].shape[0]
within_1_std / float(df_scaled.shape[0])
```

0.75

[snippet.python](#)

```
# 2
within_2_std = df_scaled[(df_scaled['friends_scaled'] <= 2) &
(df_scaled['friends_scaled'] >= -2)].shape[0]
within_2_std / float(df_scaled.shape[0])
```

0.9166666666666666

[snippet.python](#)

```
# 3
within_3_std = df_scaled[(df_scaled['friends_scaled'] <= 3) &
(df_scaled['friends_scaled'] >= -3)].shape[0]
within_3_std / float(df_scaled.shape[0])
```

1.0

Plugin Backlinks:

From:
<https://moro.kr/> - Various Ways

Permanent link:
<https://moro.kr/open/%EA%B8%B0%EC%B4%88-%ED%86%B5%EA%B3%84>

Last update: 2020/06/02 09:25



