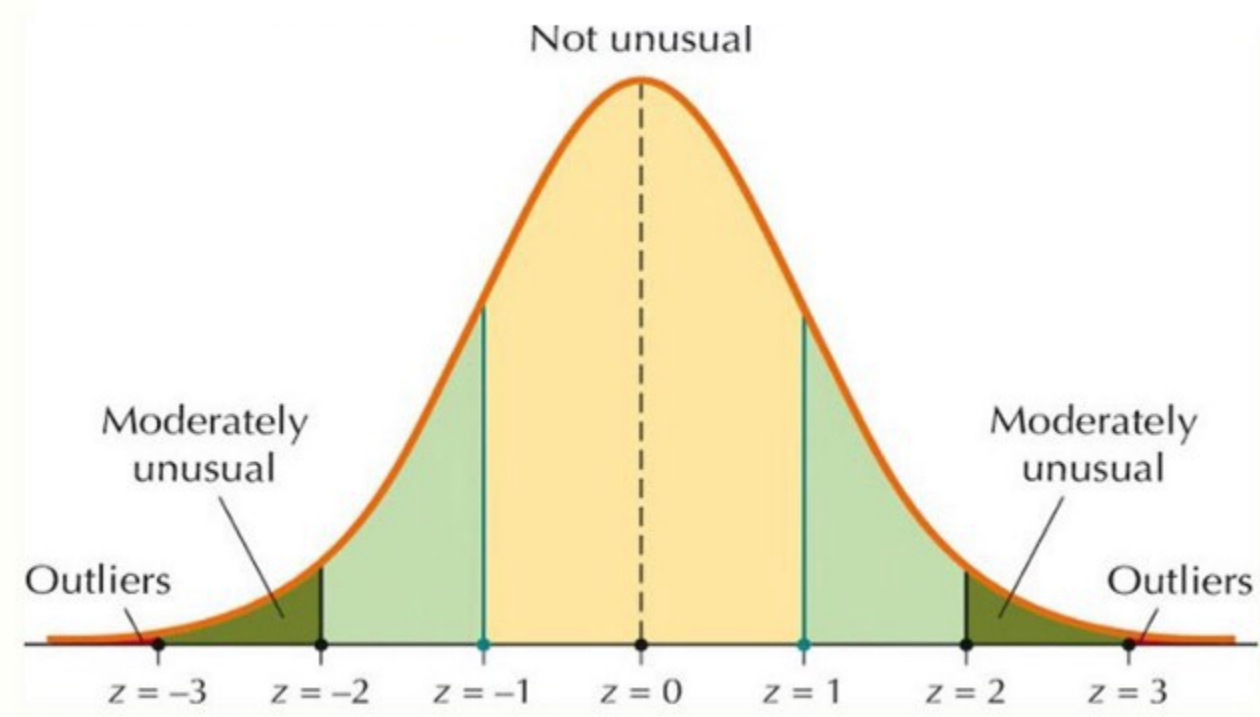


Detecting Outliers with z-Scores



```
In [1]: import numpy as np
import pandas as pd
```

```
In [2]: import pandas as pd
df=pd.read_csv(r"C:\Users\spoin\Downloads\heights.csv")
df
```

Out[2]:

	name	height
0	mohan	5.9
1	maria	5.2
2	sakib	5.1
3	tao	5.5
4	virat	4.9
5	khusbu	5.4
6	dmitry	6.2
7	selena	6.5
8	john	7.1
9	imran	14.5
10	jose	6.1
11	deepika	5.6
12	yoseph	1.2
13	binod	5.5

```
In [3]: upper_limit=df['height'].mean()+3*df['height'].std()
lower_limit=df['height'].mean()-3*df['height'].std()
print('upper_limit',upper_limit)
print('lower_limit',lower_limit)

upper_limit 14.389410604519316
lower_limit -2.2894106045193157
```

```
In [4]: # find the outlier
df[(df['height']>upper_limit) | (df['height']<lower_limit) ]
```

Out[4]:

	name	height
9	imran	14.5