# Chapter 2: Getting to know your data

| Tính mean (trung bình)    Tính mean có trọng số    Tính median (trung vị)    L1: lower boundary of the median interval  N: number of values in the data set  : sum of the frequencies of all of the intervals that are lower than the median interval  : is the frequency of the median interval  width: the width of median interval | Quartiles    - IQR: interquartile range (Khoảng cách từ Q3 tới Q1)  **Five number summary** bao gồm: Min, Q1, Q2 (Median), Q3, Max  5 số này có thể dùng để vẽ biểu đồ boxplot:   * Typically, the ends of the box are at the quartiles so that the box length is the interquartile range. * The median is marked by a line within the box. * Two lines (called whiskers) outside the box extend to the smallest (Minimum) and largest (Maximum) observations. |
| --- | --- |
| Công thức chuẩn hóa min-max:  A: Tập dữ liệu ban đầu  v’i: giá trị sau khi chuẩn hóa  vi: giá trị ban đầu (thuộc tập A)  maxA: giá trị lớn nhất trong tập A  minA: giá trị nhỏ nhất trong tập A  newMaxA: giá trị tối đa của phạm vi mới  newMinA: giá trị tối thiểu của phạm vi mới |  |

|  |  |  |
| --- | --- | --- |

|  |  |  |
| --- | --- | --- |

|  |  |  |
| --- | --- | --- |

|  |  |  |
| --- | --- | --- |