**Histograms**

 People often say that a picture is worth a thousand words. The ability to summarize a large data set is important. The three tools used most often to summarize data in Microsoft Excel are histograms, descriptive statistics, and PivotTables.

A histogram is a commonly used tool to summarize data. Essentially, a histogram tells you how many observations (another term for data points) fall in various ranges of values. For example, a histogram created from monthly Cisco stock returns might show how many monthly returns Cisco had from 0 percent through 10 percent, 11 percent through 20 percent, and so on. The ranges in which you group data are referred to as bin ranges.

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

|  |  |
| --- | --- |
| **Data Range**  | **Frequency**  |
| 0-10  | 1  |
| 10-20  | 3  |
| 20-30  | 6  |
| 30-40  | 4  |
| 40-50  | 2  |

**Note:** Changing the size of the bin changes the appearance of the graph and the conclusions you may draw from it |

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

 |

|  |
| --- |
| **Example**A teacher grades the tests and creates bins of width 10 points: . . . , 30-39, 40-49, 50-59, 60-69, 70-79, . . . . The number of test scores in each data bin is recorded and plotted as a bar graph.  |
|

|  |
| --- |
| **Data**  |
| Student  | Grade  |
| Bullwinkle  | 84  |
| Rocky  | 91  |
| Bugs  | 75  |
| Daffy  | 68  |
| Wylie  | 98  |
| Mickey  | 78  |
| Minnie  | 77  |
| Lucy  | 86  |
| Linus  | 94  |
| Asterix  | 64  |
| Obelix  | 59  |
| Donald  | 54  |
| Sam  | 89  |
| Taz  | 76  |

 | Professor Chang's Physics class has just taken a test. In order to come up with meaningful grades, Professor Chang will make a histogram to represent the distribution of grades and find a reasonable central value. The critical question is that of bin size. Clearly, a bin size of 100 makes no sense as it puts all the data in one bin, giving us no information. At the same time, a bin size of 1 or less makes no sense as the bins would be so small as to look pretty much like a simple list of results. We already have that! Let's try a few bin sizes: * bin width of 20
* bin width of 3
* bin width of 10

This makes it sound like 10 is the best width. Actually, we don't know that. First of all, 8 or 12 might be better. Secondly, narrower or wider bins might give us the look at the data that we need in a particular case.  |

|  |  |  |
| --- | --- | --- |
| http://quarknet.fnal.gov/run2/graphics/3bin.gif | http://quarknet.fnal.gov/run2/graphics/10bin.gif | http://quarknet.fnal.gov/run2/graphics/5bin.gif |

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

