# Statistics 3

## Qualitative Variables

### MA 101 Enrollment

| Category | Number of Students | Percentage | Pie Angle Size |
|---|---|---|---|
| Freshman | 57 | | |
| Sophomore | 8 | | |
| Junior | 13 | | |
| Senior | 44 | | |
| Other | 9 | | |
| Total | 131 | 100% | |

**Categories** --  The values of a qualitative variable are often called categories or classes.

 **Pie Chart --**  An alternative to the bar graph in the case of a qualitative variable with a small number of categories.

1

# Pie Chart of MA 101 Enrollment

## Turning Quantitative Variables into Qualitative Ones

Class intervals, the intervals into which the data values of a quantitative variable are subdivided, are normally all the same size. However, Prof. Blackbeard used unequal ones in assigning letter grades.

### Stat 101 Mid Term Exam Grades

| Class Intervals | Letter Grade | Frequency | Percentage |
|---|---|---|---|
| 0-8 | F | 10 | |
| 9-10 | D | 26 | |
| 11-13 | C | 30 | |
| 14-17 | B | 8 | |
| 18-25 | A | 1 | |
| Total | | 75 | 100% |

We plot the resulting relative frequencies versus the assigned letter grades. By doing so, we have converted a quantitative variable (test scores) into a qualitative variable (letter grades).

3

# Pie Chart of Stat 101 Letter Grades

# Histograms

When the quantitative variable is continuous, we use a special type of bar graph called a *histogram*. In a histogram there are no gaps between the class intervals. **Endpoint convention** – we agree that salaries that fall exactly on a boundary between class intervals are placed in the lower class interval.

**Starting salaries of UAB graduates, 1992**

| Salary Range | Nbr. Of Graduates | Percent |
|---|---|---|
| 20,000 – 25,000 | 228 | 7% |
| 25,000 – 30,000 | 456 | 14% |
| 30,000 – 35,000 | 1043 | 32% |
| 35,000 – 40,000 | 912 | 28% |
| 40,000 – 45,000 | 391 | 12% |
| 45,000 – 50,000 | 163 | 5% |
| 50,000 – 55,000 | 65 | 2% |
| Total | 3258 | 100% |

# Histogram of UAB Graduate Starting Salaries
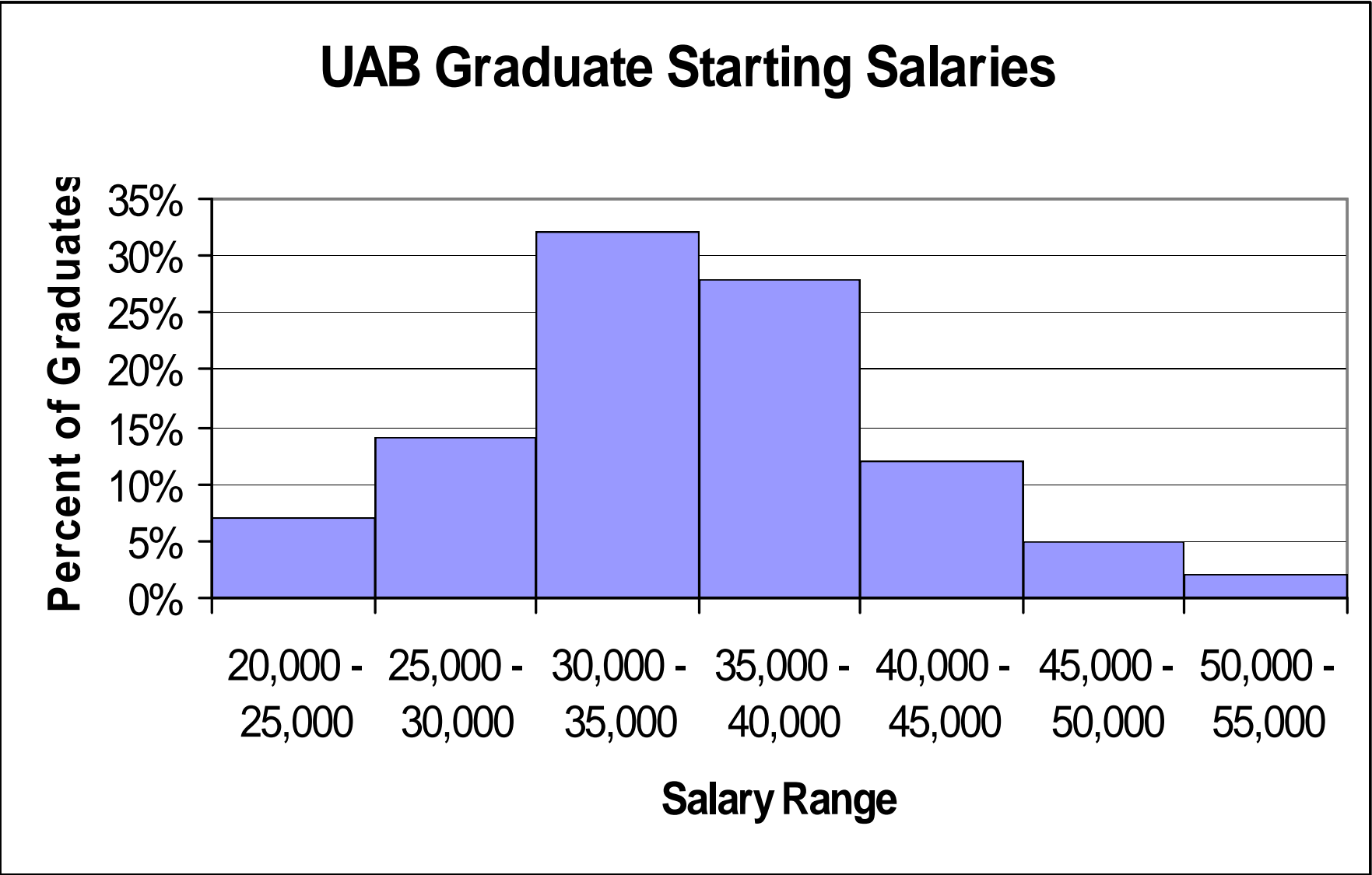
# Five Number Summary

The following five numbers provide a useful numerical summary of a of a data set, and for that reason are called collectively the *five number summary*.

## Stat 101 Exam

|  | Score |
|---|---|
| • Minimum: *Min* | 1 |
| • First Quartile: $Q_1$ | 9 |
| • Median: *M* | 11 |
| • Third Quartile: $Q_3$ | 12 |
| • Maximum: *Max* | 24 |

# Box Plots

A convenient and eye-catching way to display the five number summary is the *box plot.*

The *horizontal axis* is the scale on which the data values fall.

The *central box* runs from the first quartile $Q_1$ to the third quartile $Q_3$.

The median is represented by a *vertical line* at its appropriate location inside the central box.

The *whiskers* on each side reach from the quartiles to the Min and Max data values.

**Example.** Recall that the star forward of UAB's basketball team in his first 12 games makes the following *sorted* scores.

3   18   20   21   22   24   25   26   27   29   33   39

Min = 3      $Q_1$= 20.5      M = 24.5      $Q_3$= 28      Max = 30

Basketball Player's Scores in 12 Games

# Using Box Plots for Comparisons

*Example.* Box plots are particularly useful for driving home comparisons for two or more populations.

| SCHOOL | MIN | $Q_1$ | MEDIAN | $Q_3$ | MAX |
|---|---|---|---|---|---|
| Engineering | 28,000 | 35,000 | 40,000 | 44,000 | 49,000 |
| Humanities | 22,000 | 25,000 | 34,000 | 39,000 | 54,000 |

# Box Plots on Same Axis

## Observations

- The starting salaries in the Humanities are more spread out.

- Three-fourths of Humanities graduates earn less than the median of Engineering graduates.

-