

Unit 2: Descriptive Statistics

Marius Ionescu

09/06/2011

Definition

Definition

- To summarize data, statisticians often use a graph called a *histogram*.

Definition

- To summarize data, statisticians often use a graph called a *histogram*.
- The histogram is a bar graph with height determined by the proportion of the population.

Definition

- To summarize data, statisticians often use a graph called a *histogram*.
- The histogram is a bar graph with height determined by the proportion of the population.

Definition

- To summarize data, statisticians often use a graph called a *histogram*.
- The histogram is a bar graph with height determined by the proportion of the population.
- It tells what the distribution of the graph is in terms of the variable that is studied.

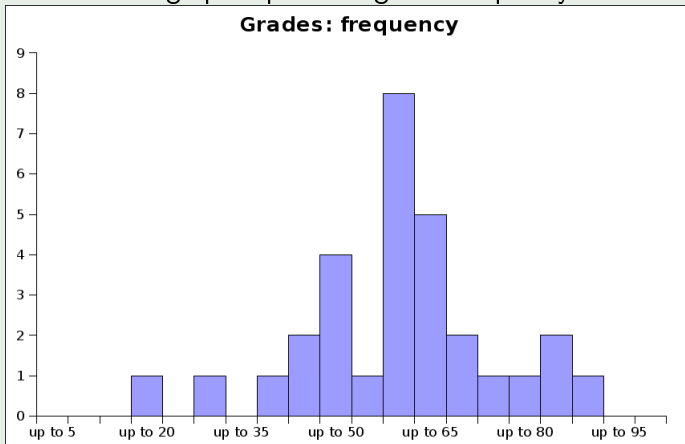
Definition

- To summarize data, statisticians often use a graph called a *histogram*.
- The histogram is a bar graph with height determined by the proportion of the population.
- It tells what the **distribution** of the graph is in terms of the variable that is studied.
- **In a histogram the area of the blocks represent percentages.**

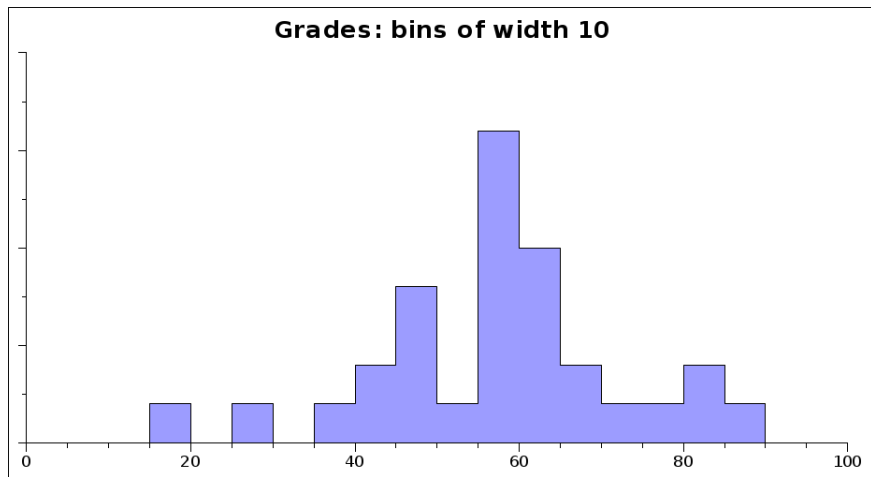
Example: Grades in a class

Example

This is a bar graph representing the “frequency”

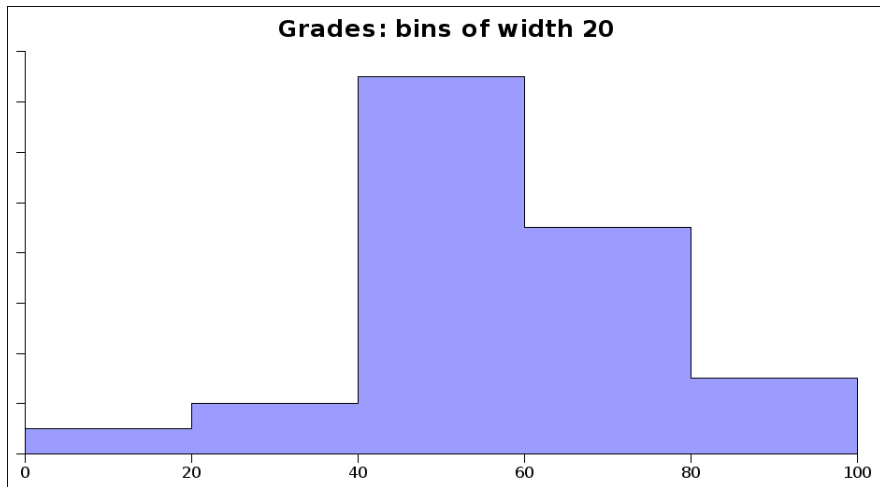


Scores: no scale on the y-axis

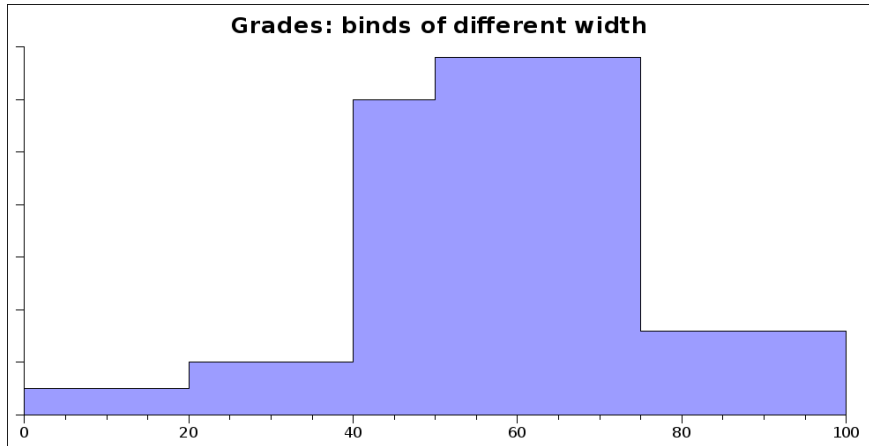


Correction: bins have width 5!

Scores: bin of width 20



Determining the grade: bins of variable width



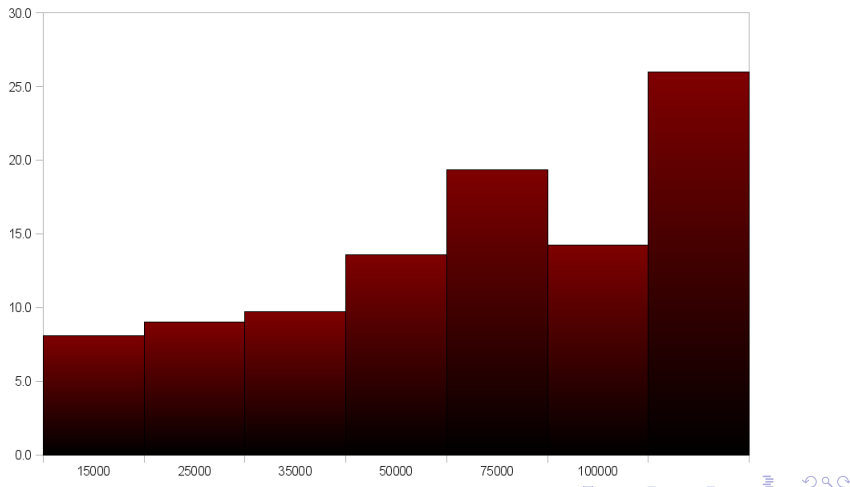
Income distribution (in thousand dollars) - 2008

<\$15	\$15-24.9	\$25-\$34.9	\$35-\$49.9	\$50-\$74.9	\$75-\$99.9	\$100
8.1	9.0	9.7	13.6	19.4	14.3	26.0

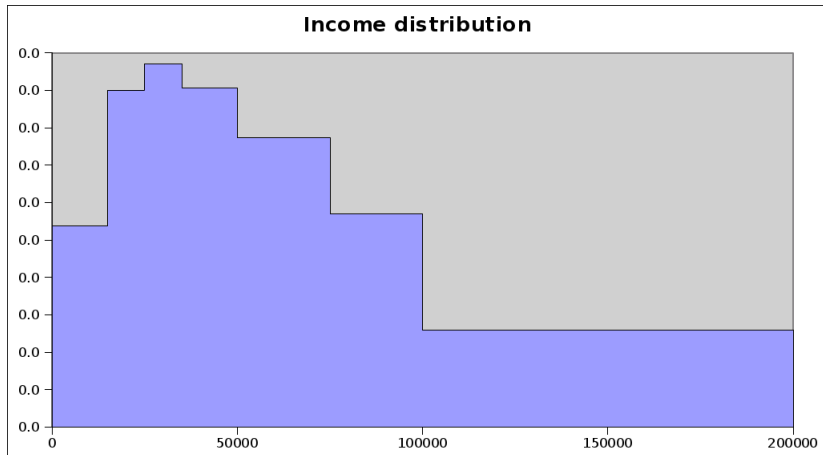
Income distribution (in thousand dollars) - 2008

<\$15	\$15-24.9	\$25-\$34.9	\$35-\$49.9	\$50-\$74.9	\$75-\$99.9	\$100
8.1	9.0	9.7	13.6	19.4	14.3	26.0

Income distribution 2008



The correct histogram for income distribution



The Density Scale

Fact

Fact

- *In a histogram the area of the blocks represent percentages.*

Fact

- *In a histogram the area of the blocks represent percentages.*
- *In a histogram, the height of a block represents crowding - percentage per horizontal unit.*

Comparison of graphs using histograms

Example

Which graph has higher blood pressure?

