3.3 – Data Displays: Categorical versus Quantitative

Data Display # 1 – Categorical

– **categorical variable (in a display)** → data that can be displayed using names or labels

Ex: 1.) blonde, brunette , red , black , etc. → category = hair (colors)

2.) red , yellow , green , purple , etc. → category = colors

3.) collie, shepherd , terrier , labrador , etc. → category = dog breeds

- **pie chart** – a type of categorical data display which uses a circle divided into sectors where each "slice" represents a portion of the whole (percents %)

- **bar graph** – a type of categorical data display which uses numerical labels that's represented by rectangles of equal width (note: between each "bar" – there are gaps) (histograms → no gaps)

Example 1: Express the following data as a pie chart and a bar chart.

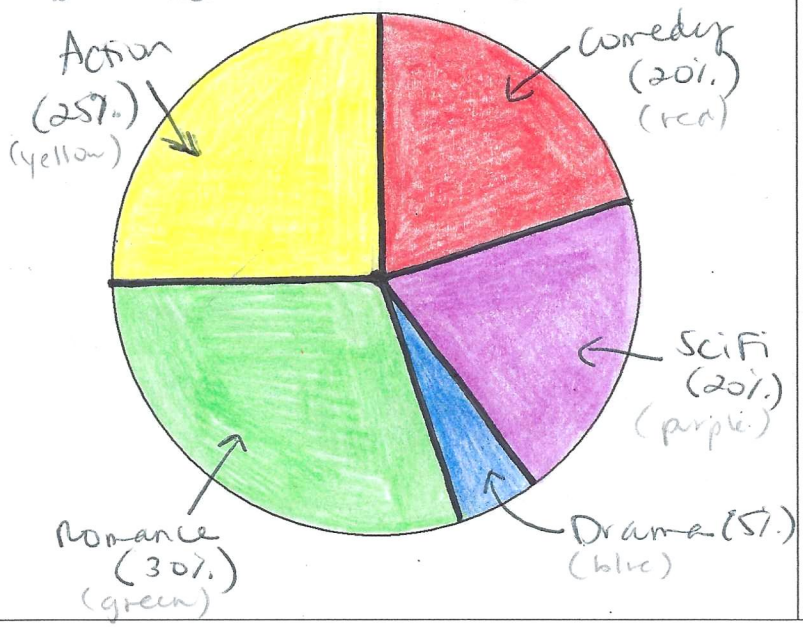
Carly did a survey with her friends about their favorite type of movies:

Class Survey: Favorite Type of Movie				
Comedy	*Action*	*Romance*	*Drama*	*SciFi*
4	5	6	1	4

Data Display # 1 – Categorical: Pie Chart

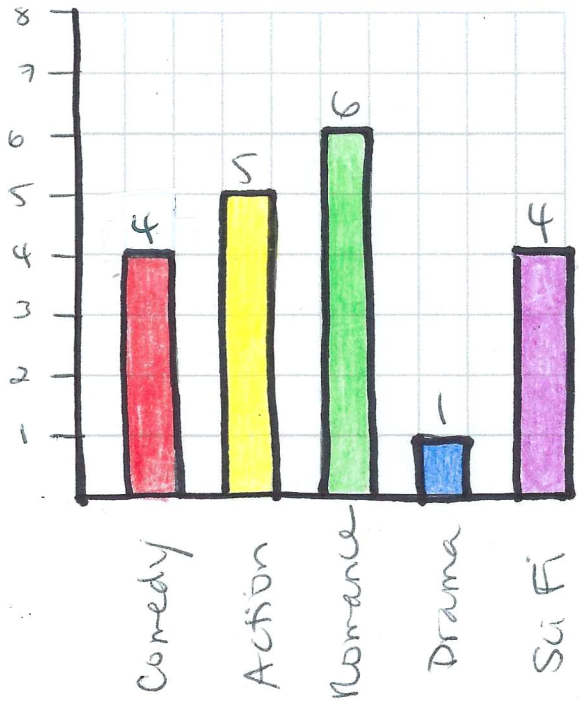
Comd	Act	Rom	Dram	SciFi	Total
4	5	6	1	4	Total # = 20
20%	25%	30%	5%	20%	Total % = 100%
72°	90°	108°	18°	72°	Total Deg = 360°

.2×360 = 72° .25×360 = 90° .3×360 = 108° .05×360 = 18° .2×360 = 72°



Data Display # 1 – Categorical: Bar Graph

Title: Favorite Type of Movie



Data Display # 2 – Quantitative

– **quantitative variable (in a display)** → data that can be displayed using numerical values

Ex: 1.) 1.2 million , 4.3 million , 2.5 million ,etc. → category = population or money

2.) 8 lb , 9 lb , 12 lb , 27 lb , etc. → category = weights

3.) 36 inches , 23 inches , 48 inches , etc. → category = heights or lengths

- **stem-and-leaf plot** – a type of quantitative data display that is organized from least → greatest and separated into 2 columns (mainly useful for organizing long list of numbers)

- **box-and-whisker plot** – a type of quantitative data display that is organized in quartiles (4 equal parts) and shows the "spread" (min/max, median, range (and outliers)) of a set of data

5 number summary = min, LQ, median, UQ, and max

Example 2: Express the following data as a stem-and-leaf plot and a box-and-whisker plot.

The class scores on a 50-item test are shown in the table below.

~~71~~	~~95~~	~~84~~	~~98~~	~~88~~	~~74~~
~~90~~	~~89~~	~~86~~	~~42~~	~~99~~	~~86~~
~~91~~	~~73~~	~~66~~	~~87~~	~~89~~	~~80~~

Data Display # 2 – Quantitative: Stem-Leaf Plot

Steam	Leaf
4	2
5	
6	6
7	1 3 4
8	0 4 6 6 7 8 9 9
9	0 1 5 8 9

86.5

Find the following:

a.) mean: 82.7 b.) median = 86.5

c.) mode: 86 and 89 d.) range = 99-42 = 57

e.) variance: 176.9 (σ^2) (13.3^2)

f.) standard deviation: 13.3 (σ)

Data Display # 2 – Quantitative: Box-Whisker Plot

minimum (min) → the lowest # that is not an outlier

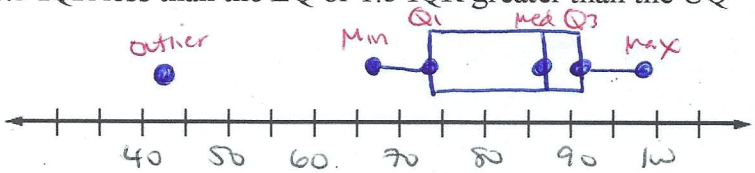
maximum (max) → the highest # that is not an outlier

lower quartile (LQ / Q1) → median of lower half of data

upper quartile (UQ / Q3) → median of upper half of data

interquartile range (IQR) → range of the middle half of data and contains 50% of data set: IQR = UQ – LQ

outlier → an element of a set of data that's at least 1.5 IQR less than the LQ or 1.5 IQR greater than the UQ



Min = 66 LQ = Q1 = 74

Max = 99 UQ = Q3 = 90

Median = 86.5 IQR = 90 - 74 = 16

low bound = 74 - 1.5(16) = 50

Any outliers? upper bound = 90 + 1.5(16) = 114

Yes → 42

*Note: LB = Q1 − 1.5(IQR)

UB = Q3 + 1.5(IQR)