

1. How many TVs do you have in your household? _____
2. How many siblings do you have? _____

TV's (x)	Siblings (y)	TV's (x)	Siblings (y)
1	2	1	5
1	3	1	5
1	3	1	2
1	2	1	2
1	2	1	2
1	2	1	2
1	2	1	2
1	2	1	2
1	2	1	2
1	2	1	2
1	2	1	2
1	2	1	2
1	2	1	2
1	2	1	2
1	2	1	2
1	2	1	2
1	2	1	2
1	2	1	2
1	2	1	2

1. How tall are you in inches? _____
2. What is your shoe size? _____

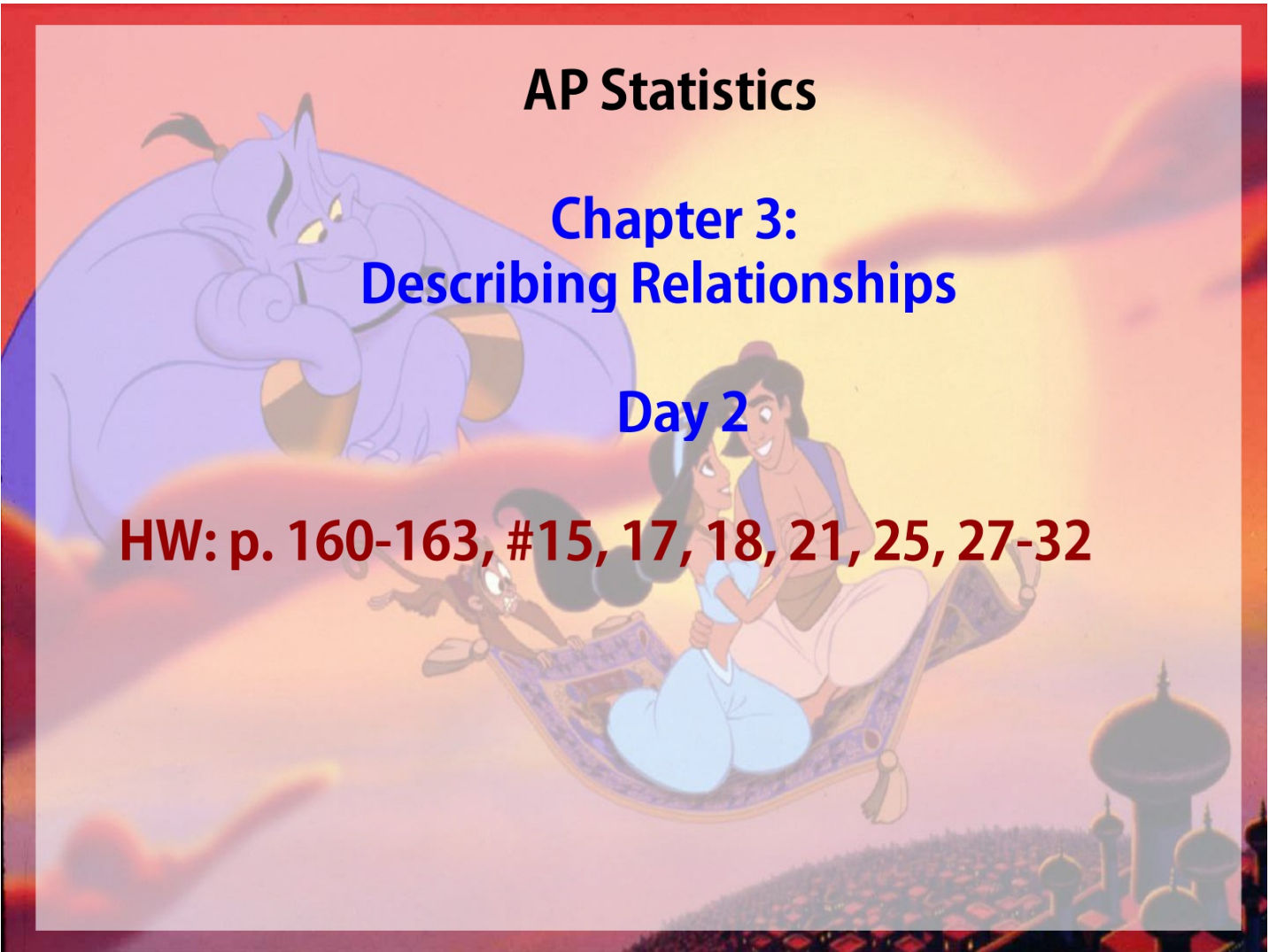
Height (x)	Shoe (y)	Height (x)	Shoe (y)
61	7	65	10
63	9	66	8.5
66.5	8	64	8.5
63	8	66	8
61	7	60	6
70.5	11	73	11
71	9	73	11
65	7	62	6.5
65	7	69	12
60	6		
60	6		
66	10		
64	10		
67	10		
63.5	10		

AP Statistics

Chapter 3: Describing Relationships

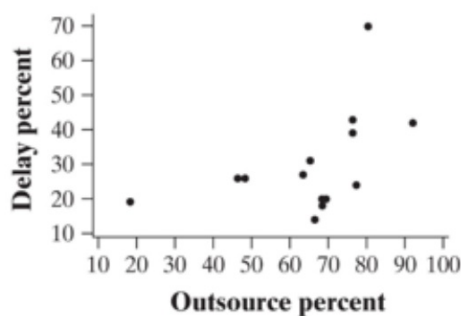
Day 2

HW: p. 160-163, #15, 17, 18, 21, 25, 27-32



p. 158-160, #1, 5, 7, 11

1. Water temperature is the explanatory variable, and weight change (growth) is the response variable. Both are quantitative.
5. A scatterplot is shown below.



7.
 - (a) The scatterplot shows a positive, somewhat curved, moderately weak association.
 - (b) The outlier is Hawaiian Airlines. Without this outlier, the relationship is more linear but still not very strong.
11.
 - (a) Most of the southern states blend in with the rest of the country. Several southern states do lie at the lower edges of their clusters. This means that, in general, the students in the southern states do not do as well as their counterparts in other portions of the country.
 - (b) West Virginia is an outlier because it has a much lower mean SAT Math score than the other states which have a similar percent of students taking the exam.

Correlation Coefficient

- Numerical measure used to judge the linear relation between two variables.
- * Correlation coefficients do not work with nonlinear relationships.
- Denoted by the letter r and is calculated using the following formula:

$$r = \frac{1}{n - 1} \sum \left(\frac{x_1 - \bar{x}}{s_x} \right) \left(\frac{y_1 - \bar{y}}{s_y} \right)$$

Chapter 3
*will not
have to
calculate by
hand.

total # of
observations

$$-1 \leq r \leq 1$$

Standard
deviations

mean

observation

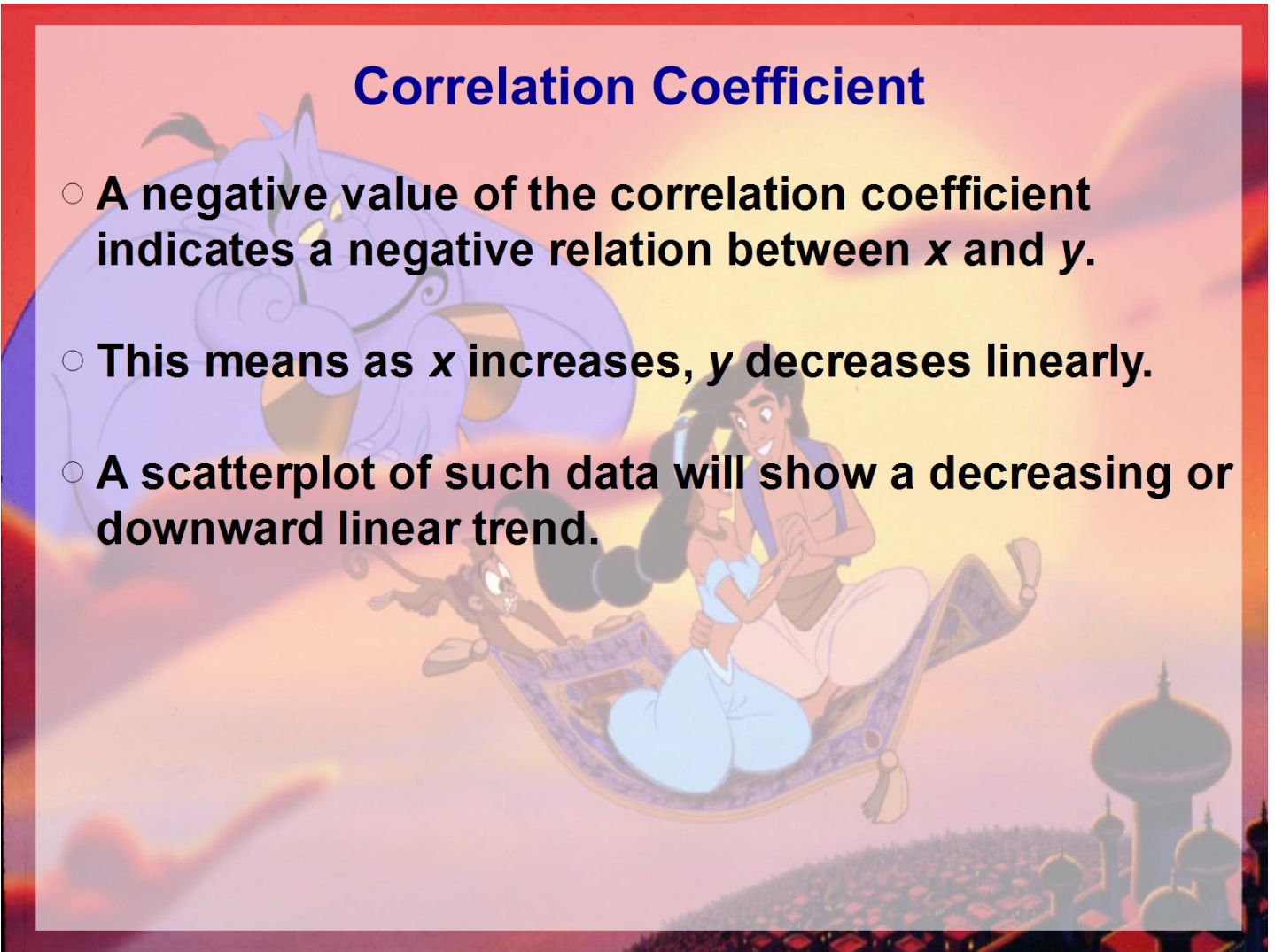
Correlation Coefficient

Direction: The positive or negative sign of the correlation describes the direction of the linear relation between the two variables.

- A positive value of the correlation coefficient indicates a positive relation between x and y .
- This means as x increases, y also increases linearly.
- A scatterplot of such data will show an increasing or upward linear trend.

Correlation Coefficient

- A negative value of the correlation coefficient indicates a negative relation between x and y .
- This means as x increases, y decreases linearly.
- A scatterplot of such data will show a decreasing or downward linear trend.



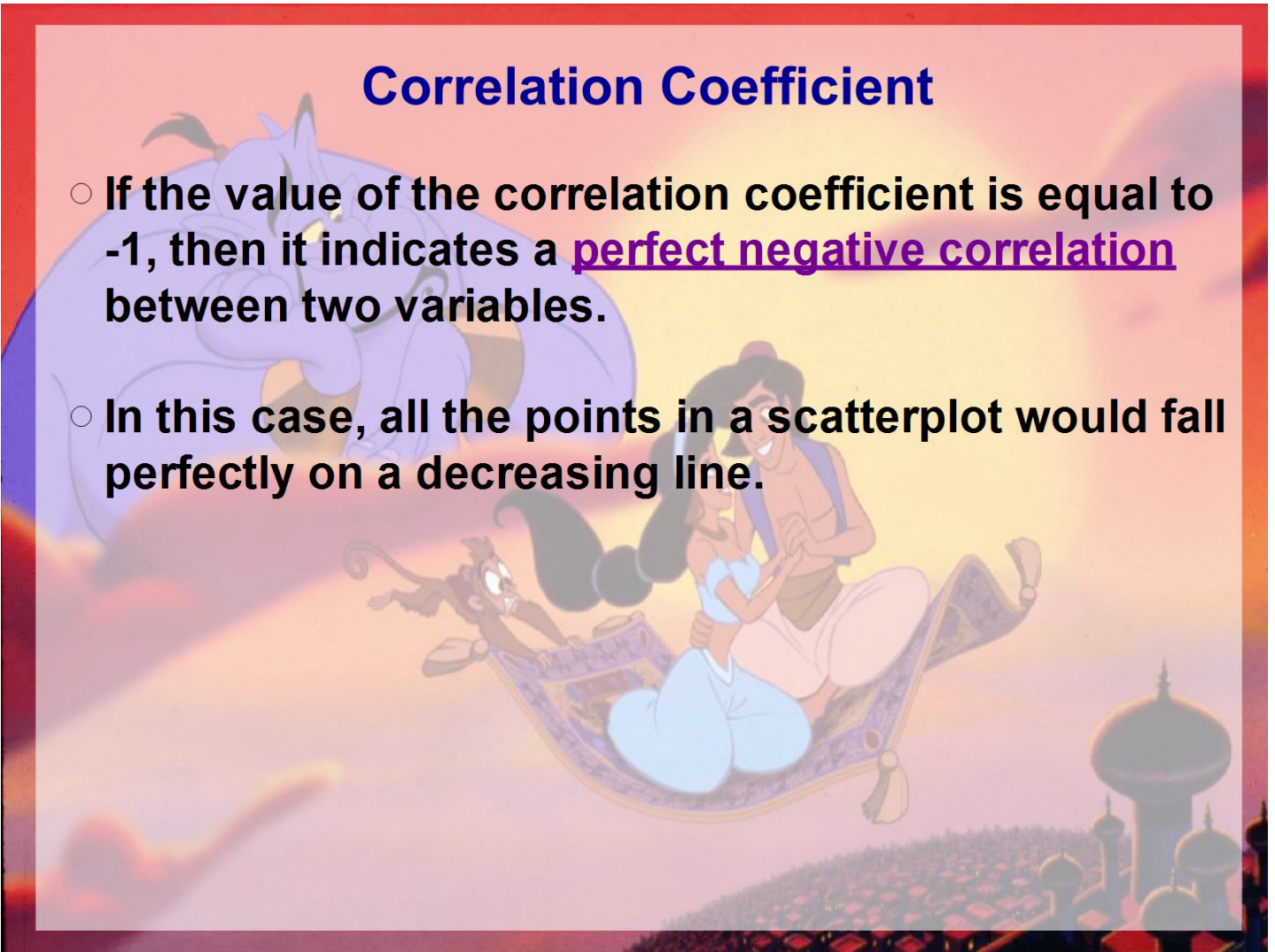
Correlation Coefficient

Strength: The numeric value of the correlation describes the strength of the linear relation between the two variables.

- If the value of the correlation coefficient is equal to +1, then it indicates a **perfect positive correlation** between two variables.
- In this case, all the points in a scatterplot would fall perfectly on an increasing line.

Correlation Coefficient

- If the value of the correlation coefficient is equal to -1 , then it indicates a perfect negative correlation between two variables.
- In this case, all the points in a scatterplot would fall perfectly on a decreasing line.



Correlation Coefficient

- If the value of the correlation coefficient is close to +1 or -1, then it indicates a strong relationship between two variables.
- If the value of the correlation coefficient is close to 0, then it indicates a weak relationship between two variables.

$r=0$ "no relationship"

Correlation Coefficient



Correlation is a unitless number that has no direct interpretation.

However, some benchmarks can help us get a feel for the strength of a linear association.

- **Ignoring the sign, correlations less than 0.5 are considered weak.**
- **Those between 0.5 and 0.8 are considered moderate.**
- **Those greater than 0.8 are considered strong.**

Match each scatterplot with the correlation that best describes the linear association between the variables.

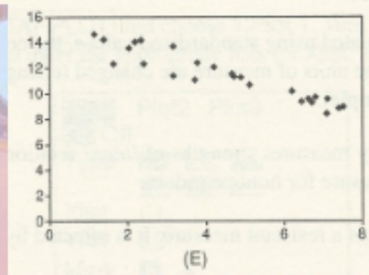
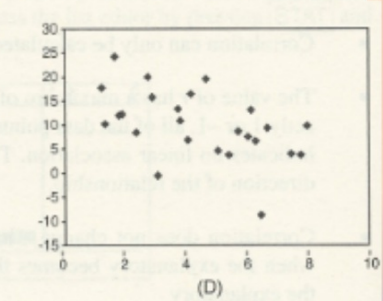
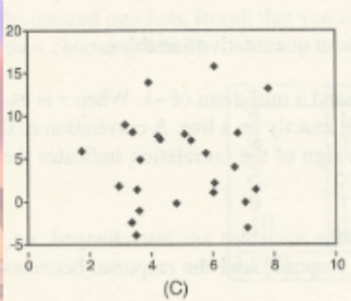
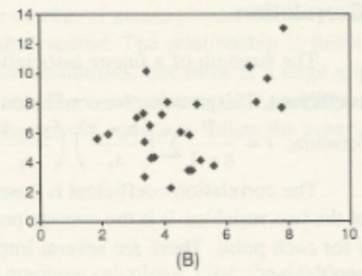
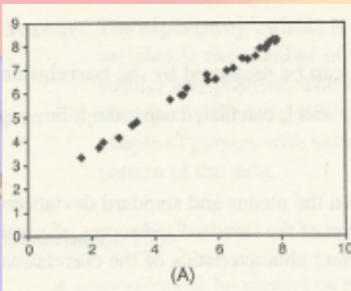
$r = 0.468$ B

$r = 0.069$ C

$r = 0.998$ A

$r = -0.940$ E

$r = -0.503$ D



Example

A sample of 12 father-and-son pairs was selected at random. The heights (in inches) of the selected pairs are below.

Height (in Inches) of Father (x)	Height (in Inches) of Son (y)
66	66
68	67
66	65
66	67
67	68
67	65
67	67
68	70
69	70
70	70
71	72
73	74

Compute the correlation coefficient and describe what it tells us about the scatterplot.

$$r = .925$$

This data is strong, positive and linear

Correlation Coefficient

Finish the "Correlation Coefficient" activity we started at the beginning of class.

