

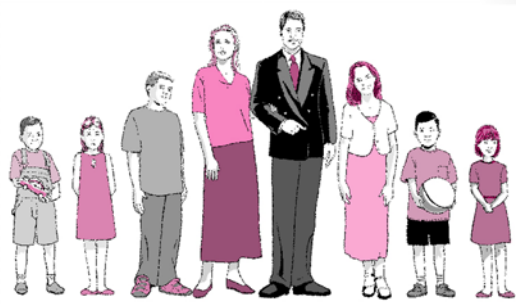
Descriptive Statistics



Estadísticas Descriptivas



Descriptive Statistics



A Normal Distribution?

Estadísticas Descriptivas



A Normal Distribution?

Estadísticas vs. Parámetros

- Un parámetro es una característica de una población.
 - Es una forma numérica o gráfica para resumir los datos obtenidos de la población
- Es una característica de una muestra.
 - Es una forma numérica o gráfica para resumir los datos obtenidos de una muestra

Statistics vs. Parameters

- A parameter is a characteristic of a population.
 - It is a numerical or graphic way to summarize data obtained from the population
- A statistic is a characteristic of a sample.
 - It is a numerical or graphic way to summarize data obtained from a sample

Tipos de Datos Numéricos

- Existen dos tipos fundamentales de datos numéricos:
 - Datos categóricos: obtenidos para determinar la frecuencia de ocurrencias en cada una de varias categorías
 - Datos cuantitativos: obtenido mediante la determinación de posición en una escala que indica la cantidad o grado

Types of Numerical Data

- There are two fundamental types of numerical data:
 - 1) Categorical data: obtained by determining the frequency of occurrences in each of several categories
 - 2) Quantitative data: obtained by determining placement on a scale that indicates amount or degree

Técnicas para Resumir Datos Cuantitativos

- Distribuciones de frecuencia
- Histogramas/ Diagrama de Tallos y Hojas
- Distribución de curvas
- Promedios/Propagación
- Variabilidad/Correlaciones

Techniques for Summarizing Quantitative Data

- Frequency Distributions
- Histograms/Stem and Leaf Plots
- Distribution curves
- Averages/Spread
- Variability/Correlations

Polígonos de Frecuencia

- Datos en algún tipo de orden
- Una distribución de frecuencias enumera de más alto a más bajo
- Esto resulta en una distribución de frecuencias agrupada. Ya que la información no es muy visual, una representación gráfica llamada polígono de frecuencia puede ayudar con esto
 - Frecuencia polígonos puede ser negativamente o positivamente sesgada
 - Pueden ser útiles para comparar dos o más grupos

Frequency Polygons

- Data in some sort of order
- A frequency distribution lists scores from high to low
- This results in a grouped frequency distribution. Since the information is not very visual, a graphical display called a frequency polygon can help with this
 - Frequency polygons can be negatively or positively skewed
 - They can be useful in comparing two or more groups

Ejemplo de una Distribución de Frecuencia

Raw Score	Frequency
64	2
63	1
61	2
59	2
56	2
52	1
51	2
38	4
36	3
34	5
31	5
29	5
27	5
25	1
24	2
21	2
17	2
15	1
6	2
3	1

n = 50

Técnicamente, la tabla debe incluir todas las puntuaciones, los que no tengan cero frecuencias incluidos. Hemos eliminado los para simplificar la presentación.

Example of a Frequency Distribution

Raw Score	Frequency
64	2
63	1
61	2
59	2
56	2
52	1
51	2
38	4
36	3
34	5
31	5
29	5
27	5
25	1
24	2
21	2
17	2
15	1
6	2
3	1

n = 50

Technically, the table should include all scores, including those for which there are zero frequencies. We have eliminated those to simplify the presentation.

Ejemplo de Distribución de Frecuencias Agrupadas

Raw Score (Intervals of Five)	Frequency
64	2
63	1
61	2
59	2
56	2
52	1
51	2
38	4
36	3
34	5
31	5
29	5
27	5
25	1
24	2
21	2
17	2
15	1
6	2
3	1

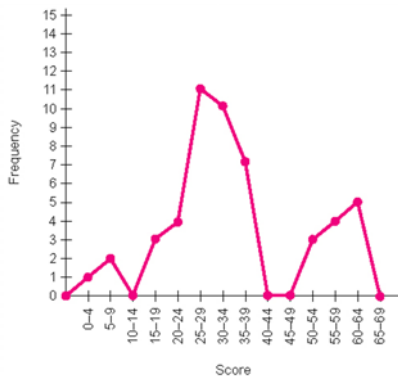
n = 50

Example of a Grouped Frequency Distribution

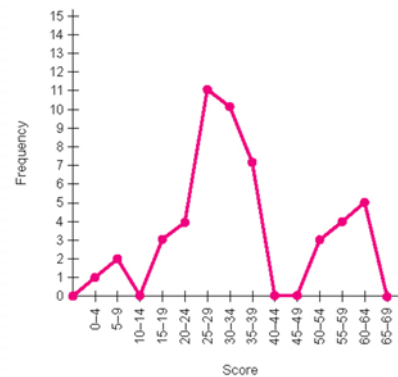
Raw Score (Intervals of Five)	Frequency
64	2
63	1
61	2
59	2
56	2
52	1
51	2
38	4
36	3
34	5
31	5
29	5
27	5
25	1
24	2
21	2
17	2
15	1
6	2
3	1

n = 50

Ejemplo de un Polígono de Frecuencia



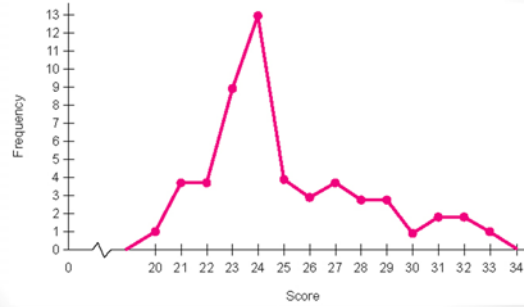
Example of a Frequency Polygon



Polígono Positivamente Sesgada



Example of a Positively Skewed Polygon



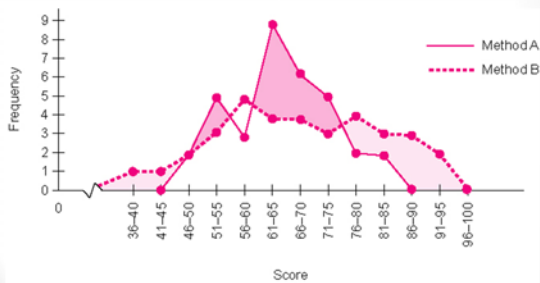
Polígono Negativamente Sesgada



Example of a Negatively Skewed Polygon



Dos Polígonos de Frecuencia en Comparación



Two Frequency Polygons Compared



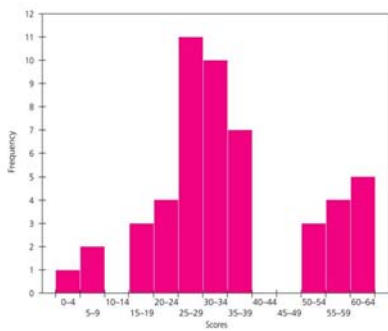
Histogramas y Diagrama de Tallo y Hojas

- Un histograma es un gráfico de barras que se utiliza para mostrar datos cuantitativos en el nivel de relación o intervalo de medición que
- Un diagrama de tallo y hojas (diagrama de tallo) parece un histograma, excepto en lugar de barras, muestra los valores para cada categoría
 - Son útiles para comparar y contrastar dos distribuciones

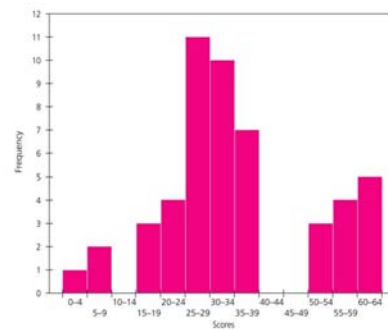
Histograms and Stem-Leaf Plots

- A histogram is a bar graph used to display quantitative data at the interval or ratio level of measurement
- A stem-leaf Plot (stem plot) looks like a histogram, except instead of bars, it shows values for each category
 - They are helpful for comparing and contrasting two distributions

Histograma de Datos



Histogram of Data



La Curva Normal

- Esta curva de distribución muestra una distribución generalizada de las puntuaciones vs. líneas rectas (polígono de frecuencia)
- Distribución de datos tiende a seguir una forma específica que se llama una distribución normal
- Esta distribución se considera "en forma de campana" y permite el trazado de los promedios siguientes:
 - Media
 - Mediana
 - Modo

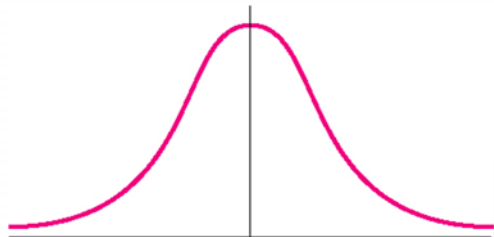
*Estas medidas de tendencia central permiten resumir los datos en una distribución de frecuencia con un único número

The Normal Curve

- This distribution curve shows a generalized distribution of scores vs. straight lines (frequency polygon)
- Distribution of data tends to follow a specific shape called a normal distribution
- This distribution is considered "bell shaped" and allows the plotting of the following averages:
 - Mean
 - Medium
 - Mode

*These measures of central tendencies enable one to summarize the data in a frequency distribution with a single number

Curva Normal



The Normal Curve



Ejemplo del Modo, Mediana y Media en una Distribución

Raw Score	Frequency
98	1
97	1
91	2
85	1
80	5
77	7
72	5
65	3
64	7
62	10
58	3
45	2
33	1
11	1
5	1
<hr/> n = 50	

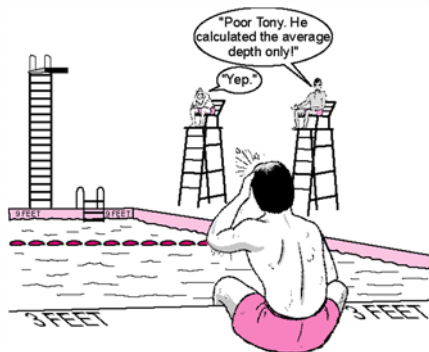
mode = 62; median = 64.5; mean = 66.7

Example of the Mode, Median and Mean in a Distribution

Raw Score	Frequency
98	1
97	1
91	2
85	1
80	5
77	7
72	5
65	3
64	7
62	10
58	3
45	2
33	1
11	1
5	1
<hr/> n = 50	

mode = 62; median = 64.5; mean = 66.7

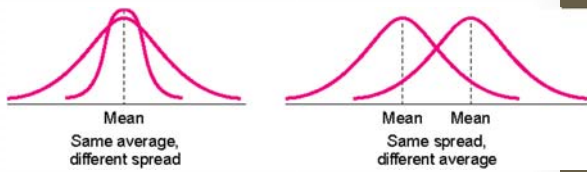
Promedios Pueden ser Engañosos



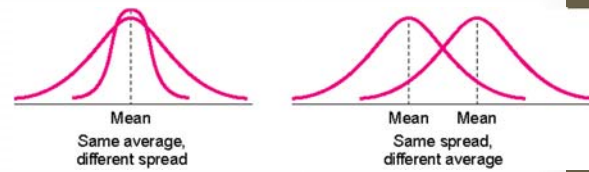
Averages Can Be Misleading



Comparación de Distribuciones Diferentes



Different Distributions Compared



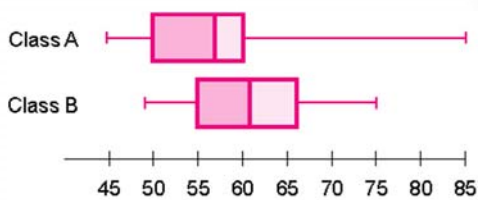
Variabilidad

- Se refiere a la medida a la que se extienden las puntuaciones de una variable cuantitativa en una distribución.
- El range representa la diferencia entre los puntajes más altos y más bajos en una distribución.
- Un resumen de cinco números informa la puntuación de más baja, el primer cuartil, la mediana, el tercer cuartil y puntuación más alta.
- Los resúmenes de cinco son a menudo retratados gráficamente mediante el uso de diagramas de caja.

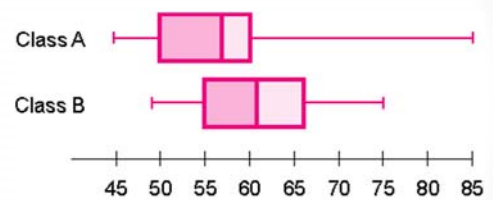
Variability

- Refers to the extent to which the scores on a quantitative variable in a distribution are spread out.
- The range represents the difference between the highest and lowest scores in a distribution.
- A five number summary reports the lowest, the first quartile, the median, the third quartile, and highest score.
- Five number summaries are often portrayed graphically by the use of box plots.

Parcelas de Caja



Box plots



Desviación Estándar

- Considerado el mejor índice de variabilidad.
- Es un número único que representa la propagación de una distribución.
- Si una distribución es normal entonces la media más o menos 3 SD abarcará alrededor del 99% de todas las puntuaciones en la distribución.

Standard Deviation

- Considered the most useful index of variability.
- It is a single number that represents the spread of a distribution.
- If a distribution is normal, then the mean plus or minus 3 SD will encompass about 99% of all scores in the distribution.

Cálculo de la Desviación Estándar de una Distribución

Raw Score	Mean	$X - \bar{X}$	$(X - \bar{X})^2$
85	54	31	961
80	54	26	676
70	54	16	256
60	54	6	36
55	54	1	1
50	54	-4	16
45	54	-9	81
40	54	-14	196
30	54	-24	576
25	54	-29	841

$$\text{Variance (SD}^2) = \frac{\sum(X - \bar{X})^2}{n}$$

$$= \frac{3640}{10} = 364^a$$

$$\text{Standard deviation (SD)} = \sqrt{\frac{\sum(X - \bar{X})^2}{n}}$$

$\sum \sigma \sqrt{X}$
 $\sum \sigma \sqrt{X}$

Calculation of the Standard Deviation of a Distribution

Raw Score	Mean	$X - \bar{X}$	$(X - \bar{X})^2$
85	54	31	961
80	54	26	676
70	54	16	256
60	54	6	36
55	54	1	1
50	54	-4	16
45	54	-9	81
40	54	-14	196
30	54	-24	576
25	54	-29	841

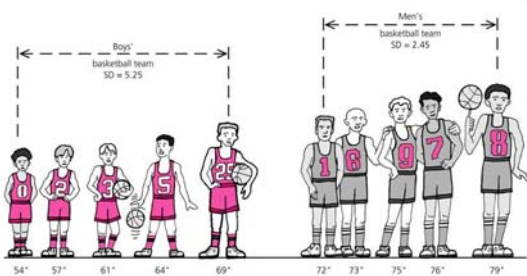
$$\text{Variance (SD}^2) = \frac{\sum(X - \bar{X})^2}{n}$$

$$= \frac{3640}{10} = 364^a$$

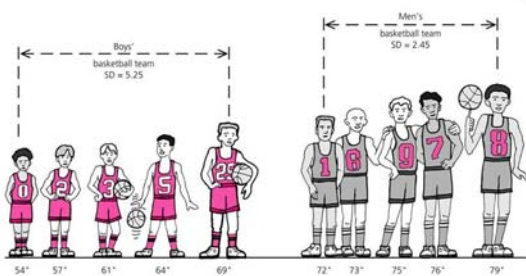
$$\text{Standard deviation (SD)} = \sqrt{\frac{\sum(X - \bar{X})^2}{n}}$$

$\sum \sigma \sqrt{X}$
 $\sum \sigma \sqrt{X}$

Desviaciones Estándar para Equipos de Baloncesto de Hombre y Jóvenes



Standard Deviations for Boys' and Men's Basketball Teams



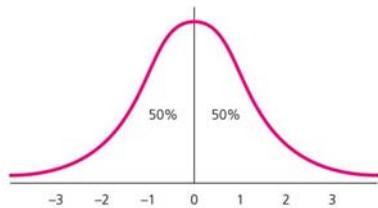
Hechos Acerca de la Distribución Normal

- 55% de todas las observaciones caen a cada lado de la media.
- 68% de las puntuaciones entran en 1 SD de la media en una distribución normal.
- 27% de las observaciones se sitúan entre 1 y 2 SD de la media.
- 99,7% de todas las puntuaciones corresponden dentro 3 SD de la media.
- Esto a menudo se conoce como la regla **68-95-99,7**

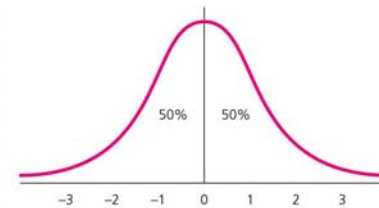
Facts about the Normal Distribution

- 55% of all the observations fall on each side of the mean.
- 68% of scores fall within 1 SD of the mean in a normal distribution.
- 27% of the observations fall between 1 and 2 SD from the mean.
- 99.7% of all scores fall within 3 SD of the mean.
- This is often referred to as the **68-95-99.7 rule**

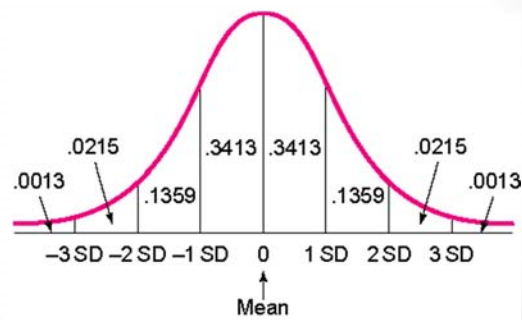
50% de todas las puntuaciones caem en una curva Normal a cada lado de la Media



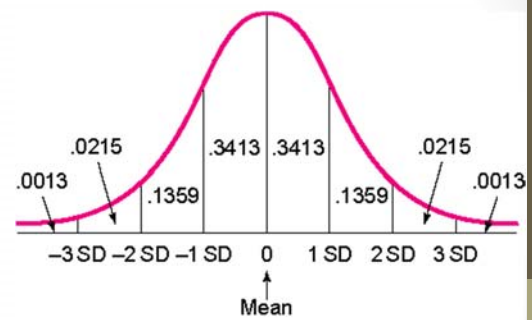
Fifty Percent of All Scores in a Normal Curve Fall on Each Side of the Mean



Probabilidades Bajo la Curva Normal



Probabilities Under the Normal Curve



Puntuaciones Estándar

- Las puntuaciones estándar utilizan una escala común para indicar cómo una persona se compara con otros individuos en un grupo.
- La forma más simple de una puntuación estándar es una puntuación Z.
- Una puntuación Z expresa qué tan lejos se encuentra una puntuación cruda de la media en unidades de desviación estándar.

Standard Scores

- Standard scores use a common scale to indicate how an individual compares to other individuals in a group.
- The simplest form of a standard score is a Z score.
- A Z score expresses how far a raw score is from the mean in standard deviation units.

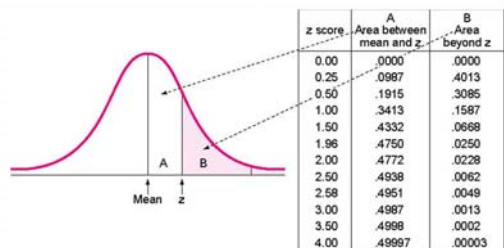
Puntuaciones Estándar

- Las puntuaciones estándar proporcionan una mejor base para comparar el rendimiento sobre las diferentes medidas que los puntajes crudos.
- Una probabilidad es un por ciento declarado en forma decimal y se refiere a la probabilidad de que se produzca un evento.
- Puntuaciones T son puntuaciones z expresados en una forma diferente (z score x 10 + 50).

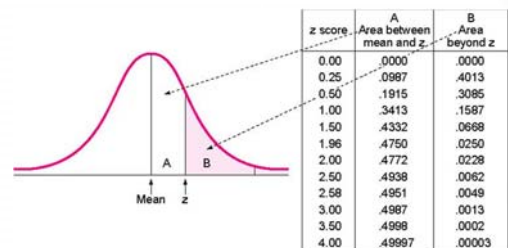
Standard Scores

- Standard scores provide a better basis for comparing performance on different measures than do raw scores.
- A Probability is a percent stated in decimal form and refers to the likelihood of an event occurring.
- T scores are z scores expressed in a different form (z score x 10 + 50).

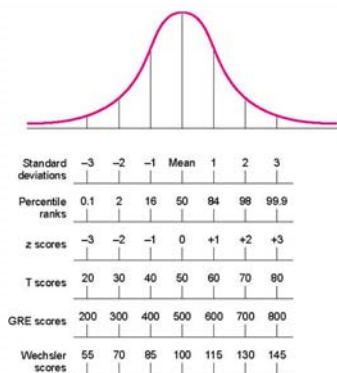
Áreas de Probabilidad entre la Media y Puntuaciones Diferentes



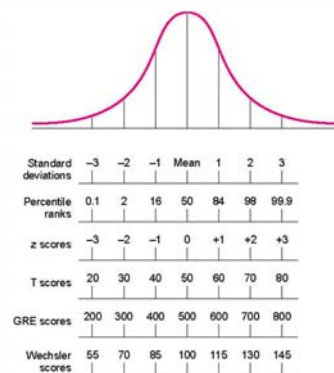
Probability Areas Between the Mean and Different Scores



Ejemplos de Puntuaciones Estándar



Examples of Standard Scores



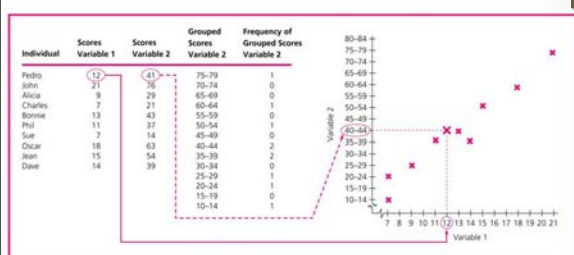
Correlación

- Los investigadores intentan determinar si existe una relación entre dos o más variables cuantitativas.
- Un **diagrama de dispersión** es una representación gráfica de la relación entre dos variables cuantitativas.
- Valores extremos** son puntuaciones que difieren o superan considerablemente la mayoría de las otras puntuaciones en una distribución o patrón.
 - Indican una excepción inusual para un patrón general
- Coefficientes de correlación** expresan el grado de relación entre dos conjuntos de resultados.
 - Coefficiente de correlación de momento del producto Pearson
 - Eta

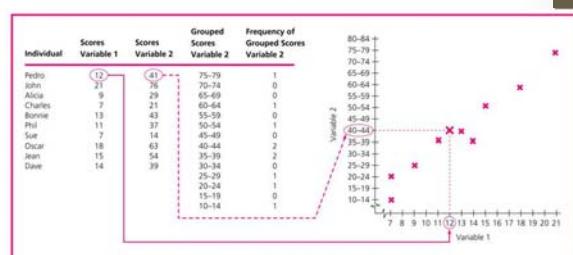
Correlation

- Researchers seek to determine whether a relationship exists between two or more quantitative variables.
- A **scatterplot** is a pictorial representation of the relationship between two quantitative variables.
- Outliers** are scores that deviate or fall considerably outside most of the other scores in a distribution or pattern.
 - They indicate an unusual exception to a general pattern
- Correlation coefficients** express the degree of relationship between two sets of scores.
 - Pearson Product-Moment Correlation Coefficient
 - Eta

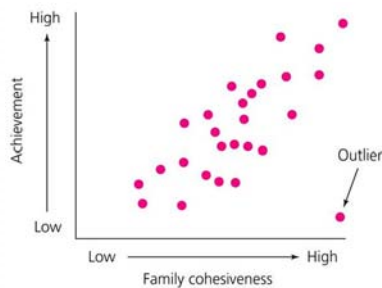
Diagrama de Dispersión de Datos



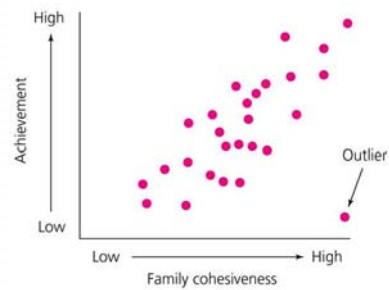
Scatterplot of Data



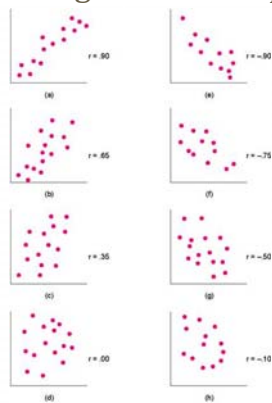
Relación entre la Cohesión de la Familia y Logros en la Escuela en un Grupo Hipotético de Estudiantes



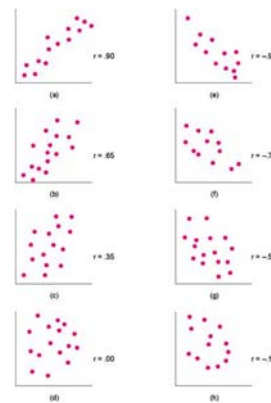
Relationship Between Family Cohesiveness and School Achievement in a Hypothetical Group of Students



Ejemplos de Diagramas de Dispersión



Examples of Scatterplots



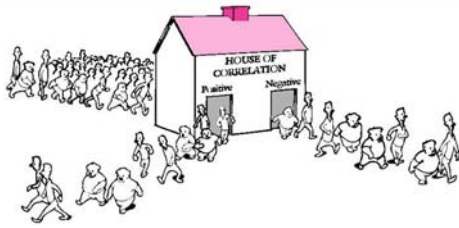
Una Correlación Negativa Perfecta



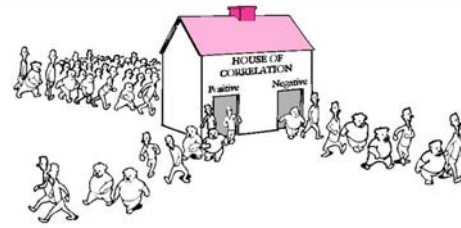
A Perfect Negative Correlation



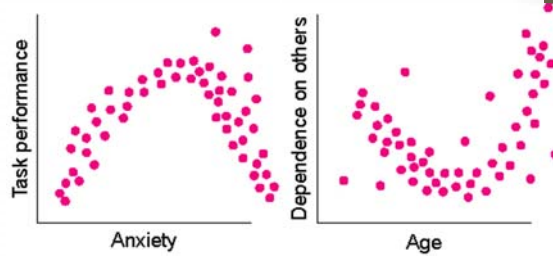
Correlaciones Positivas y Negativas



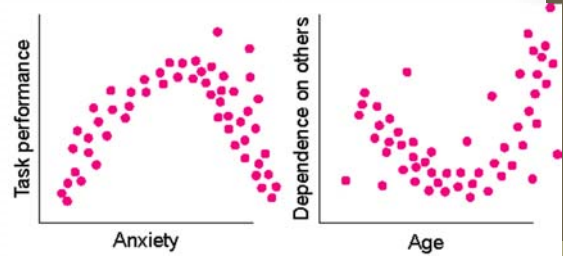
Positive and Negative Correlations



Ejemplos de Relaciones no Lineales (curvilíneas)



Examples of Nonlinear (Curvilinear) Relationships



Técnicas para Resumir Datos Categóricos

- Tabla de frecuencia
- Los gráficos de barras y gráficos circulares
- Tabla de Crossbreak

Techniques for Summarizing Categorical Data

- The Frequency Table
- Bar Graphs and Pie Charts
- The Crossbreak Table

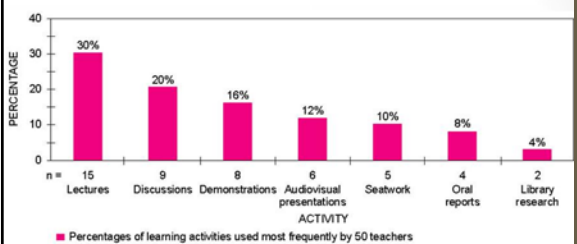
Frecuencia y Porcentaje de Respuestas al Cuestionario

Respuesta	Frecuencia	Porcentaje del Total(%)
Lección	15	30
Discusiones en clase	10	20
Demostraciones	8	16
Presentaciones audiovisuales	6	12
Trabajos	5	10
Reportes Orales	4	8
Investigación de biblioteca	2	4
Total	50	100

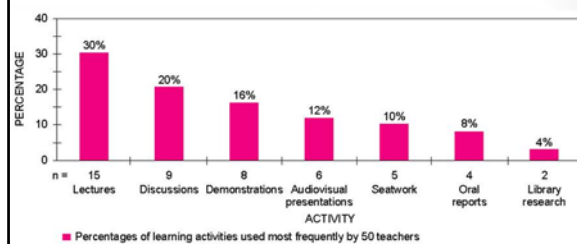
Frequency and Percentage of Responses to Questionnaire

Response	Frequency	Percentage of Total (%)
Lecture	15	30
Class discussions	10	20
Demonstrations	8	16
Audiovisual presentations	6	12
Seatwork	5	10
Oral reports	4	8
Library research	2	4
Total	50	100

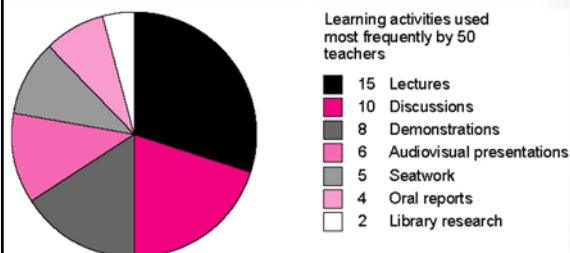
Ejemplo de una Gráfica de Barras



Example of a Bar Graph



Ejemplo de Gráfica Circular



Example of Pie Chart

