$y = f_x(x)$ SPS $a, b \in \mathbb{R}$

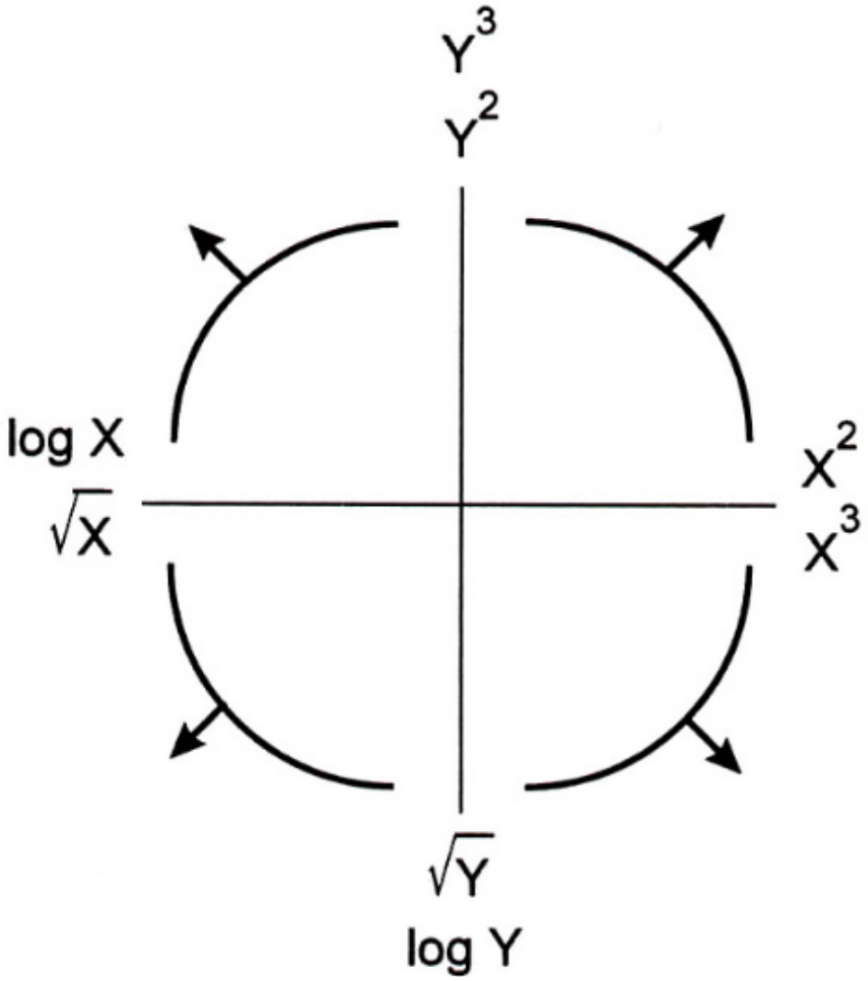
y is Linear in x then $y = ax + b$

$\log y$ $\log y = ax + b$

$y = Ce^{ax}$

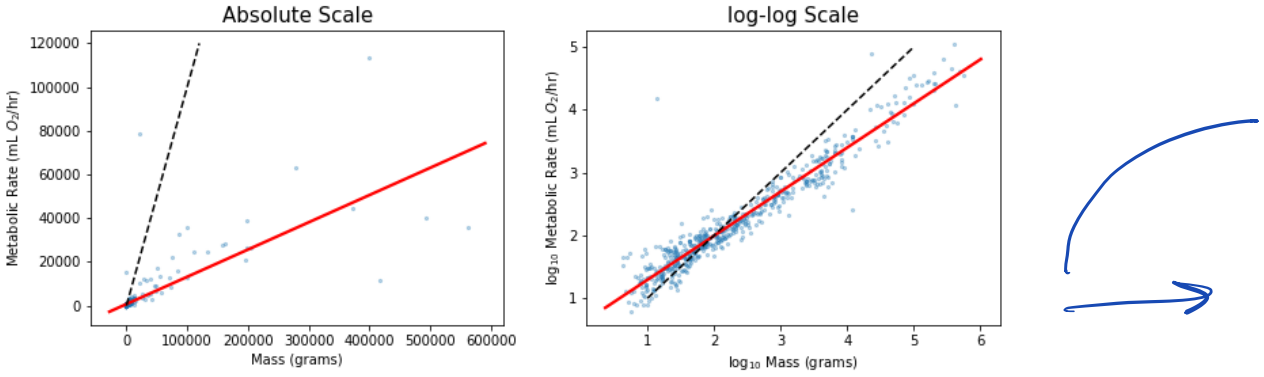
$\log y$ $\log x$ $\log y = a \log x + b$

$y = Cx^a$



Logarithmic Transformations

4. One of your friends at a biology lab asks you to help them analyze `panTHERIA`, a database of mammals. They are interested in the relationship between mass, measured in grams, and metabolic rate ("energy expenditure"), measured by oxygen use per hour. Originally, they show you the data on a linear (absolute) scale, shown on the left. You notice that the values on both axes vary over a large range with many data points clustered around the smaller values, so you suggest that they instead plot the data on a log-log scale, shown on the right. The solid red line is a "line of best fit" (we'll formalize this later in the course) while the black dashed line represents the identity line $y = x$.



(a) Let C and k be some constants and x and y represent mass and metabolic rate, respectively. Based on the plots, which of the following best describe the pattern seen in the data? Reminder: $log(ab) = log(a) + log(b)$.

$\log y = k(\log x) + C$

$e^{\log y} = e^{k \log x + C}$

$y = e^{k \log x} e^{C}$

$y = C \cdot x^k$

○ A. $y = C + kx$ ○ B. $y = C \times 10^{kx}$ ○ C. $y = C + k \log_{10}(x)$ ● D. $y = Cx^k$

(b) What parts of the plots could you use to make initial guesses on C and k?

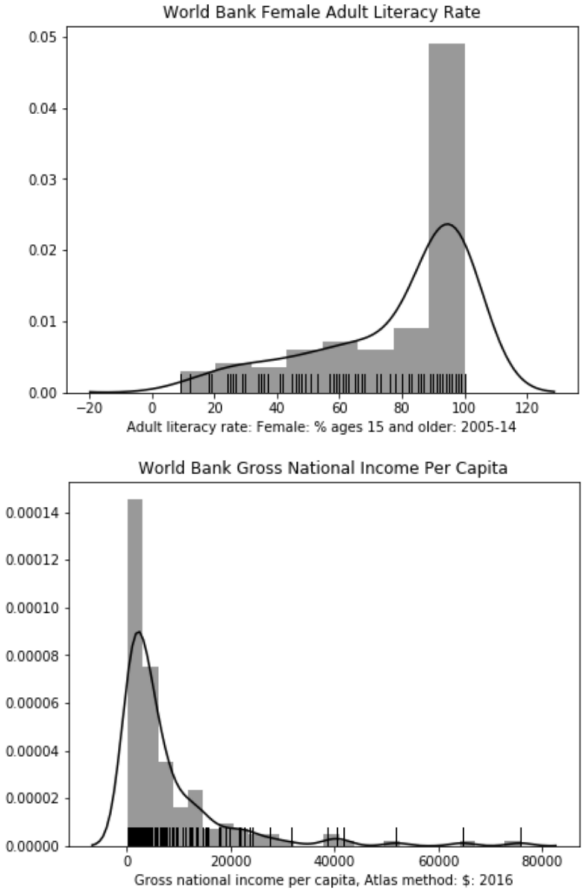
@ log-log scale, C is the y-intercept & k is the slope

(c) Your friend points to the solid line on the log-log plot and says "since this line is going up and to the right, we can say that, in general, the bigger a mammal is, the greater its metabolic rate". Is this a reasonable interpretation of the plot?

yes ✓

Visualization

3. Suppose we have the following histograms.



(a) Which of the histograms above is right skewed? **There is only one correct answer.**

◯ A. World Bank Female Adult Literacy Rate

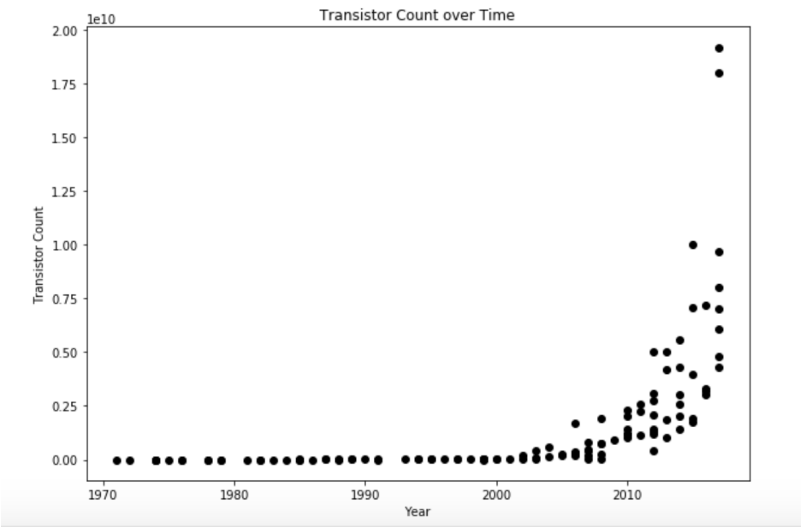
◯ B. World Bank Gross National Income Per Capita

(b) Which histogram would look more symmetric with a log transformation applied? **There is only one correct answer.**

◯ A. World Bank Female Adult Literacy Rate

◯ B. World Bank Gross National Income Per Capita

4. Consider the scatter plot shown below.



$y = e^x$

Which of the following transformations would make the relationship between x and y more linear, i.e. if we plotted $f_y(y)$ vs. $f_x(x)$, which would look most linear? **There is only one correct answer.**

- ○ A. $f_x(x) = x^2$ $\quad f_y(y) = y^2$
- ○ B. $f_x(x) = log(x)$ $\quad f_y(y) = y$
- ● C. $f_x(x) = x$ $\quad f_y(y) = log(y)$
- ○ D. $f_x(x) = log(x)$ $\quad f_y(y) = y^2$

$e^{\log y} = e^{ax+b}$

$y = e^{ax+b}$

$= e^{ax} e^{b}$

$y = C e^{ax}$

5. For each of the following relationships between x and y, select the appropriate transformation so that the transformed values are linearly related. In other words, select the transformations such that if we plotted $f_y(y)$ vs. $f_x(x)$, we'd expect to get a straight line. **There is only one correct answer in each part.**

$y = ab^x$

(a) $y = ab^x$

- ○ A. $f_y(y) = log(y)$ $\quad f_x(x) = log(x)$
- ○ B. $f_y(y) = y$ $\quad f_x(x) = log(x)$
- ● C. $f_y(y) = log(y)$ $\quad f_x(x) = x$
- ○ D. $f_y(y) = \frac{1}{y}$ $\quad f_x(x) = \frac{1}{x}$
- ○ E. The relationship is already linear.

$\log y = \log(ab^x)$

$= x \log(ab)$

$\log y = C \cdot x + 0$

(b) $y = \frac{x}{a+bx}$

- ○ A. $f_y(y) = \frac{1}{y}$ $\quad f_x(x) = x$
- ○ B. $f_y(y) = y$ $\quad f_x(x) = \frac{1}{x}$
- ○ C. $f_y(y) = \frac{1}{y}$ $\quad f_x(x) = \frac{1}{x}$
- ○ D. None of the transformations above create a linear relationship.
- ○ E. The relationship is already linear.

$$y = \frac{x}{a+bx}$$

$f(y) = \frac{1}{y}$

$$\frac{1}{y} = \frac{a+bx}{x}$$

$$= \frac{a}{x} + \frac{bx}{x}$$

$$= \frac{a}{x} + b$$

$$= \left(\frac{1}{x}\right) a + b$$

$$f(y) = a \cdot f(x) + b$$

$$\frac{1}{y} = a\left(\frac{1}{x}\right) + b$$

$f(x) = 1/x$

$$\frac{1}{y} = a \cdot f(x) + b$$

$$f(y) = a \cdot f(x) + b$$

$f(y) = 1/y \quad f(x) = 1/x$