



DATA SCIENCE

Dr. Karthik Chandrasekhar

Department of Science and Humanities

karthikchandrasekhar@pes.edu

DATA SCIENCE

Types of Transformations

Karthik Chandrasekhar

Department of Science and Humanities

DATA SCIENCE

Types of Transformations



When we see a plot of y vs x that does not admit a linear model, we transform the data in one of the many ways as follows, including but not limited to...

- $\ln y$ vs x
- $\ln y$ vs $\ln x$
- y^b vs x^a
- etc

DATA SCIENCE

Types of Transformations – In more generality



The earlier mentioned transformations can also be generalized to for a more exhaustive list of transformations that can be used.

- $(\ln y)^a$ vs x^b
- $(\ln y)^a$ vs $(\ln x)^b$
- y^b vs x^a
- etc

DATA SCIENCE

Types of Transformations – In more generality



- When seeing a heteroscedastic residual plot, using $\ln y$ or y^a instead of y where $0 < a < 1$ can be useful and is worth trying!
- When you see a trend or curve in the residual plot using y^2 vs x or y vs x^{-1} can be useful and is worth trying.
- There is however no formula that can determine which transformation to use.



THANK YOU

Karthik Chandrasekhar
Department of Science and Humanities
karthikchandrasekhar@pes.edu