

Regression and Assumptions

Background on Regression

What is Regression?

- Statistical test for prediction
- Type of modeling
- Allows for continuous IVs and continuous or categorical DVs

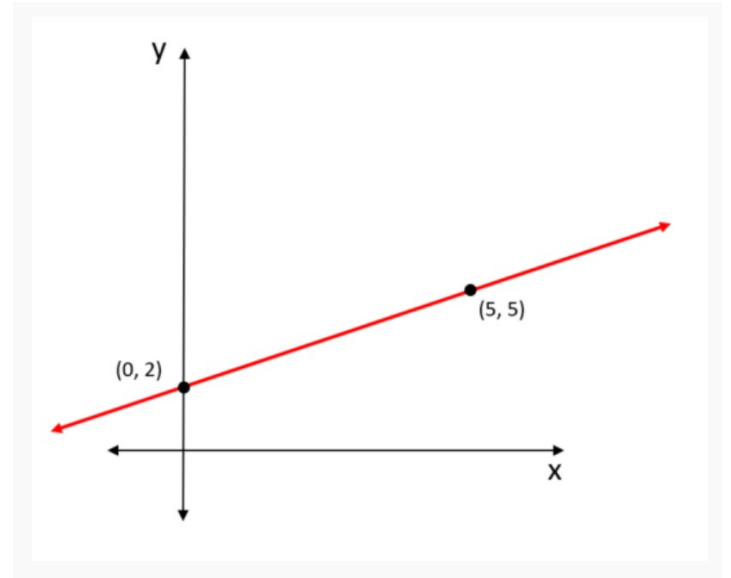
Analysis	DV Type	Levels of DV
Binary Logistic Regression	Categorical	2
Multinomial Logistic Regression	Categorical	3+
Linear Regression	Continuous	NA

Types of Regression

- Linear – Data forms a straight line
- Non-linear – Data does not form a straight line
- Logistic – DV is categorical
- Simple – One IV
- Multiple – Multiple IVs
 - Stepwise/Hierarchical – Assess each IV one at a time to see how important it is to the outcome

Parts of a Regression Line

- $y = mx + b$
- m - Slope – how steep the line is
- b - Intercept – where the line starts on the y axis
- Residual – how different the predicted line is from the actual line
 - AKA error



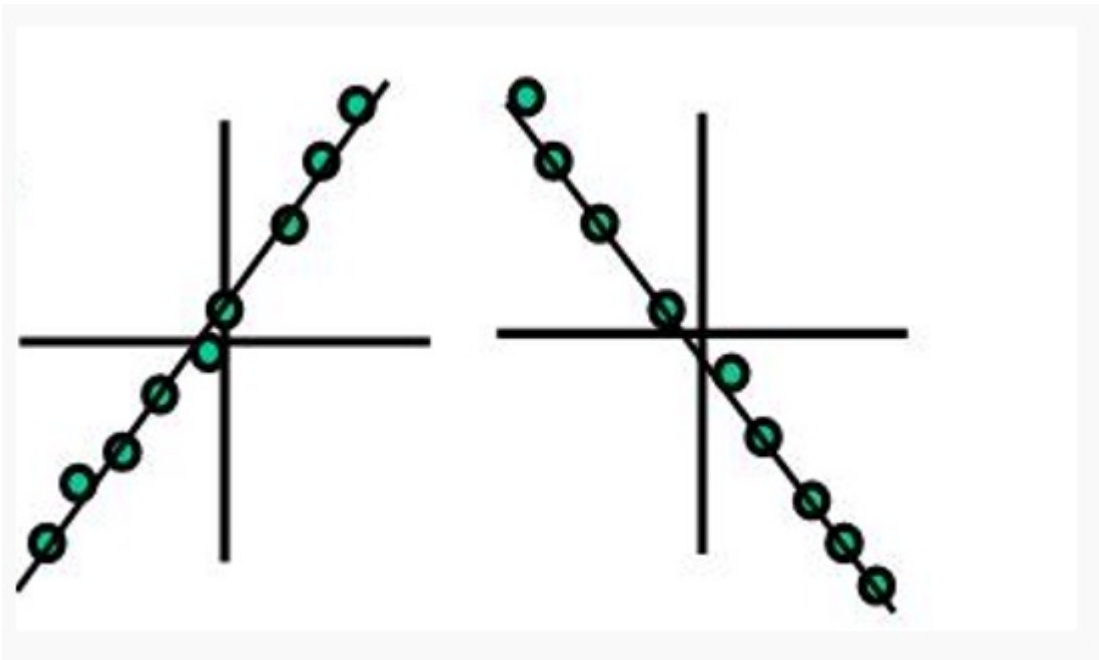
Assumptions of Linear Regression

Assumptions of Linear Regression

- Linear relationship between x and y
- Homoscedasticity – normal distribution of your error
- Homogeneity of variance – change in error is the same for all values of x
- X s are measured without error
- Absence of Multicollinearity – observations are independent without too much overlap
- Lack of outliers

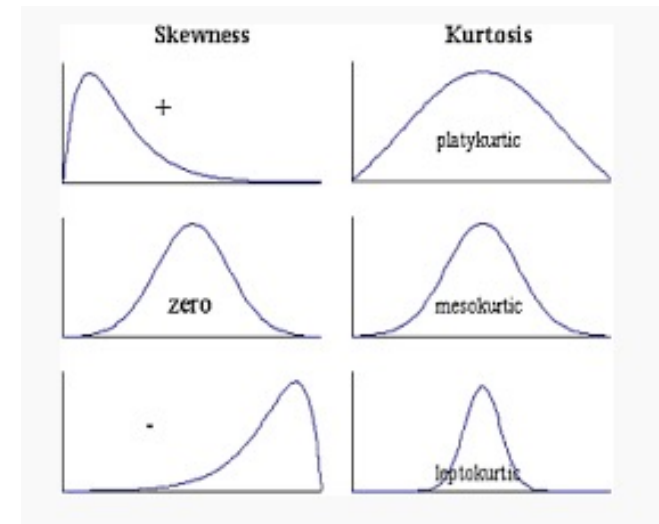
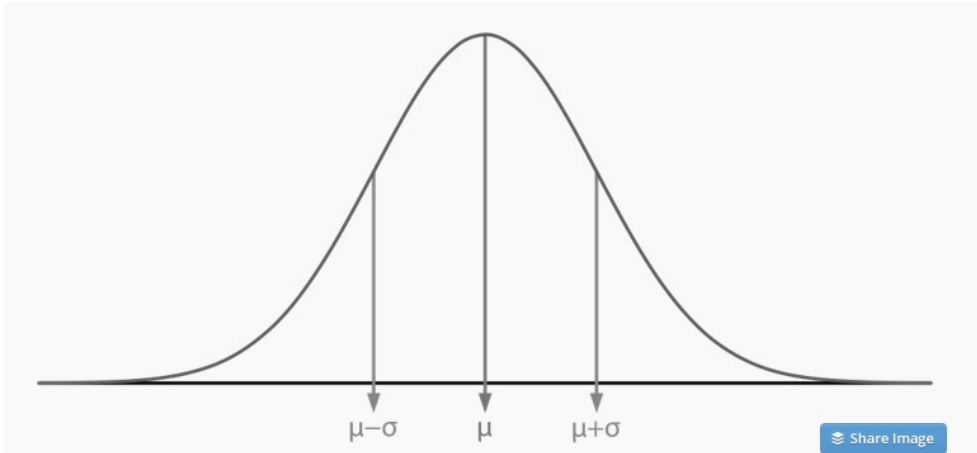
Linear

- Relationship between x and y must be linear



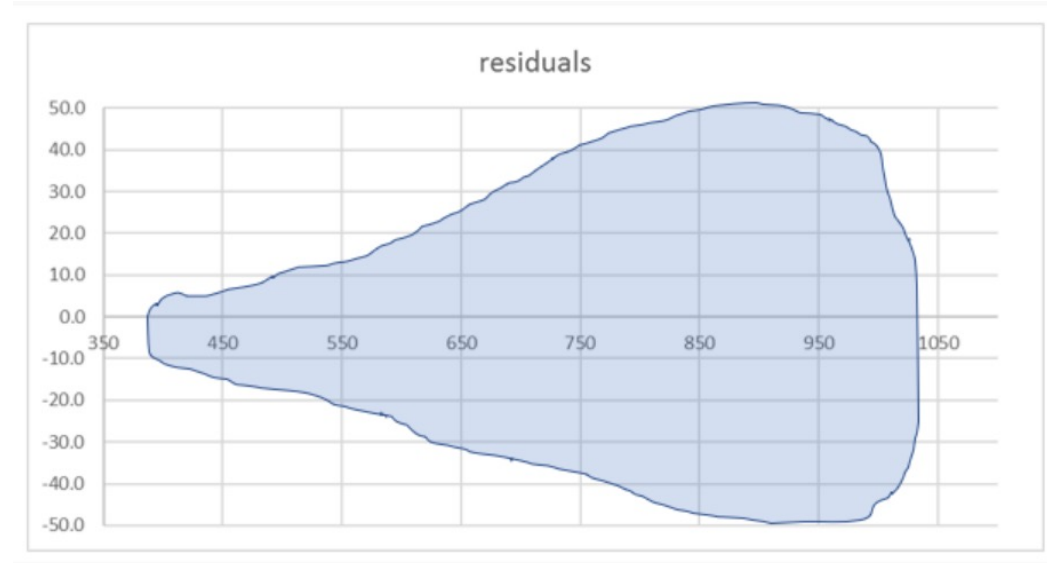
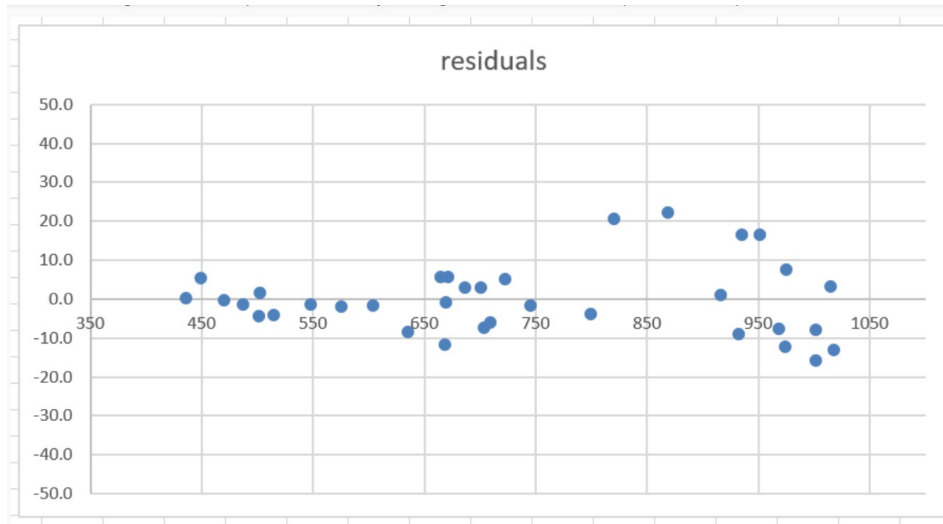
Homoscedasticity

- Normal Distribution of the error



Homogeneity of Variance

- The change in the error is the same at all levels of x



X as a Known Constant

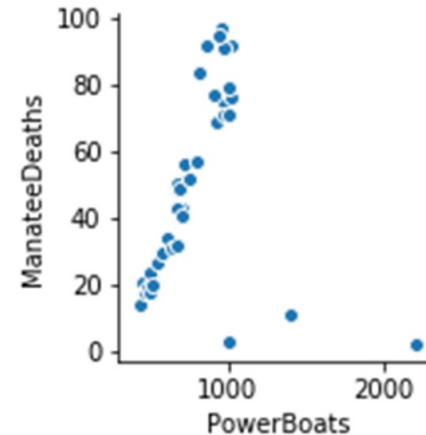
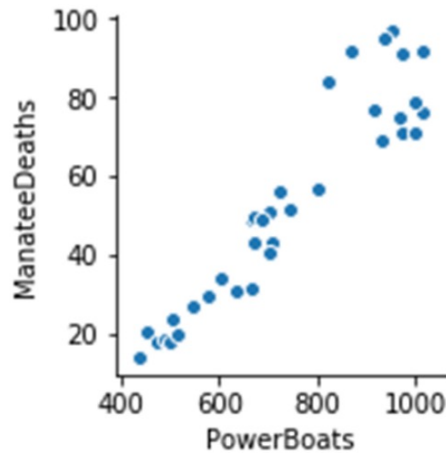
- Xs are measured without error
- Good luck verifying this or achieving it!

Absence of Multicollinearity

- Y values are independent
- Knowing something about one data point DOES NOT mean you know something about another one

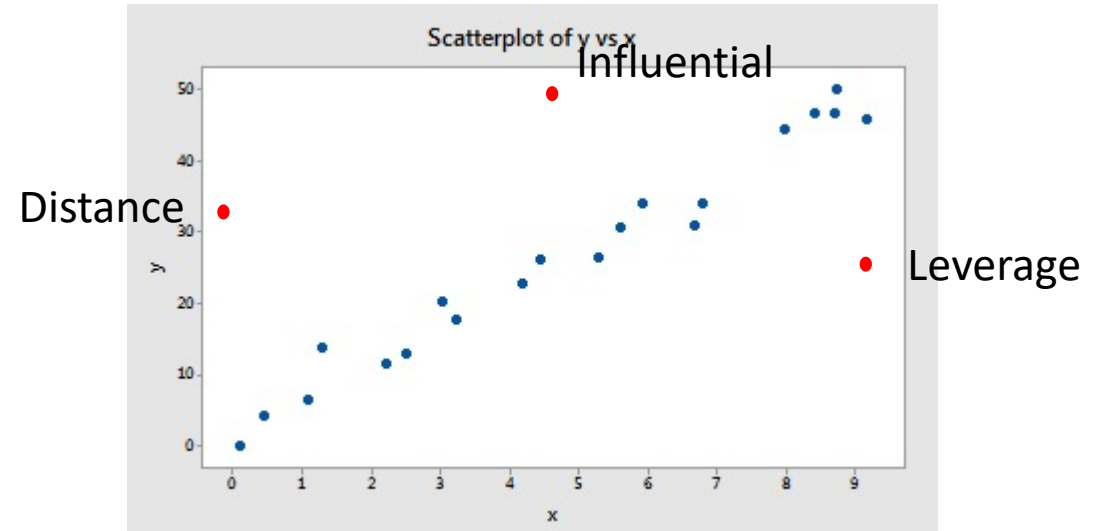
Absence of Outliers

- Data points not near any others
- Can skew your data



Types of Outliers

- Leverage - extreme value in x space
- Distance - extreme value in y space
- Influential - extreme in both x and y space



Questions?