Anscombe's Quartet: Fallacies with Descriptive Measures

DESCRIPTIVE MEASURES

Beware of Descriptive Measures

> We need to understand our data better before we make conclusions.

> We can use descriptive measures to help us understand our data.

W

DESCRIPTIVE MEASURES

Beware of Descriptive Measures

We have 6 measures of a dataset.	
Property	**Value**
mean(y)	9
var(x)	11
mean(y)	7.50
var(y)	4.125
corrcoef(x,y)	0.816
LinearRegression	y = 3 + 0.5x

What does the dataset look like?
Do the data have outliers?
Do the data form a linear relationship?
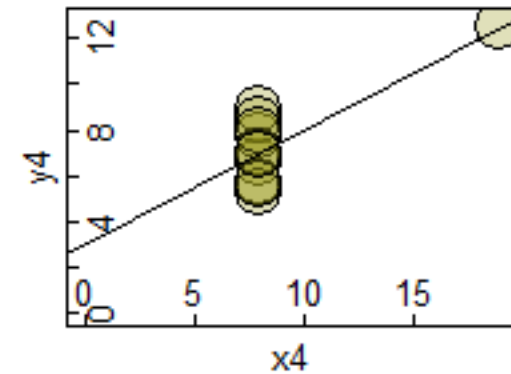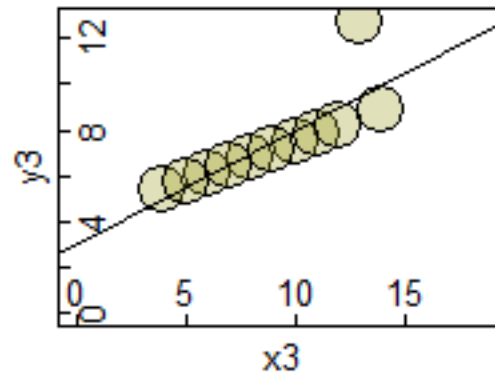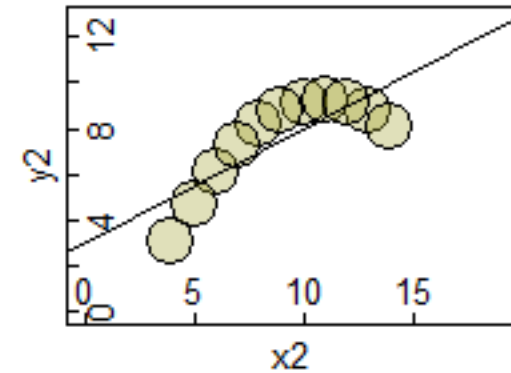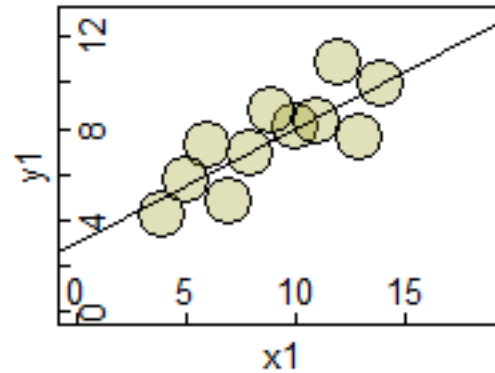Can we extrapolate from this relationship?

W

DESCRIPTIVE MEASURES

Beware of Descriptive Measures
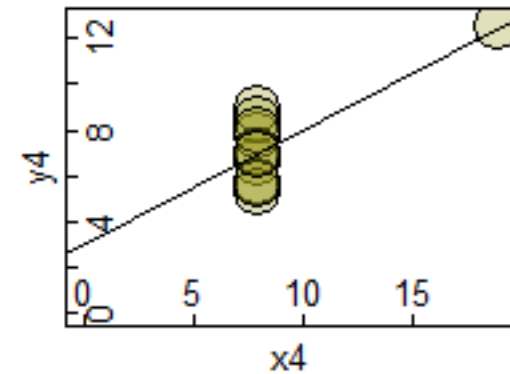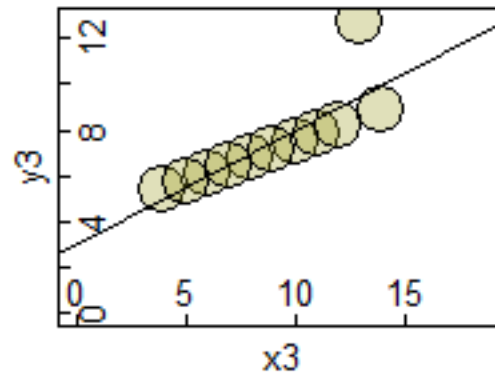
We have 6 measures of a dataset.

Property	Value
mean(y)	9
var(x)	11
mean(y)	7.50
var(y)	4.125
corrcoef(x,y)	0.816
LinearRegression	y = 3 + 0.5x

What does the dataset look like?
Do the data have outliers?
Do the data form a linear relationship?
Can we extrapolate from this relationship?

DESCRIPTIVE MEASURES

Anscombe's Quartet
See: AnscombeQuartet.R
https://en.wikipedia.org/wiki/Anscombe%27s_quartet

Beware of Descriptive Measures

We have 6 measures of a dataset.

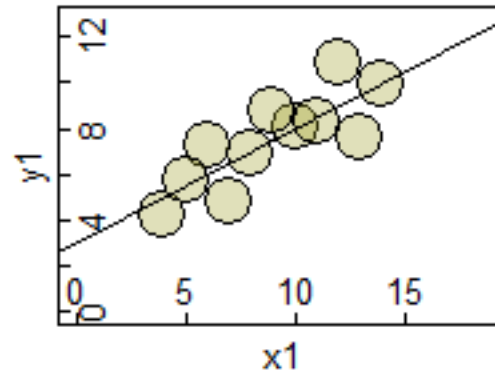| Property	Value
mean(y)	9
var(x)	11
mean(y)	7.50
var(y)	4.125
corrcoef(x,y)	0.816
LinearRegression	y = 3 + 0.5x

All these data sets have these measurements!

Anscombe's Quartet: Fallacies with Descriptive Measures