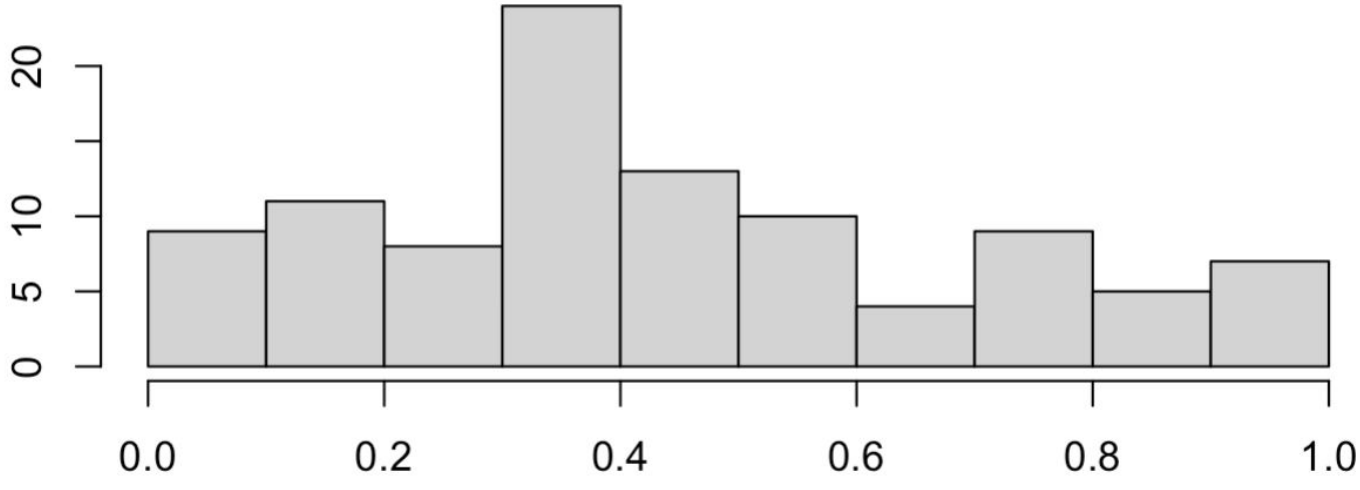
Central Limit Theorem

Pranav Madhukar
Mentored by Ronan Perry

Town X

What is my friend's political preference?

$\mathbb{E}[Y]$

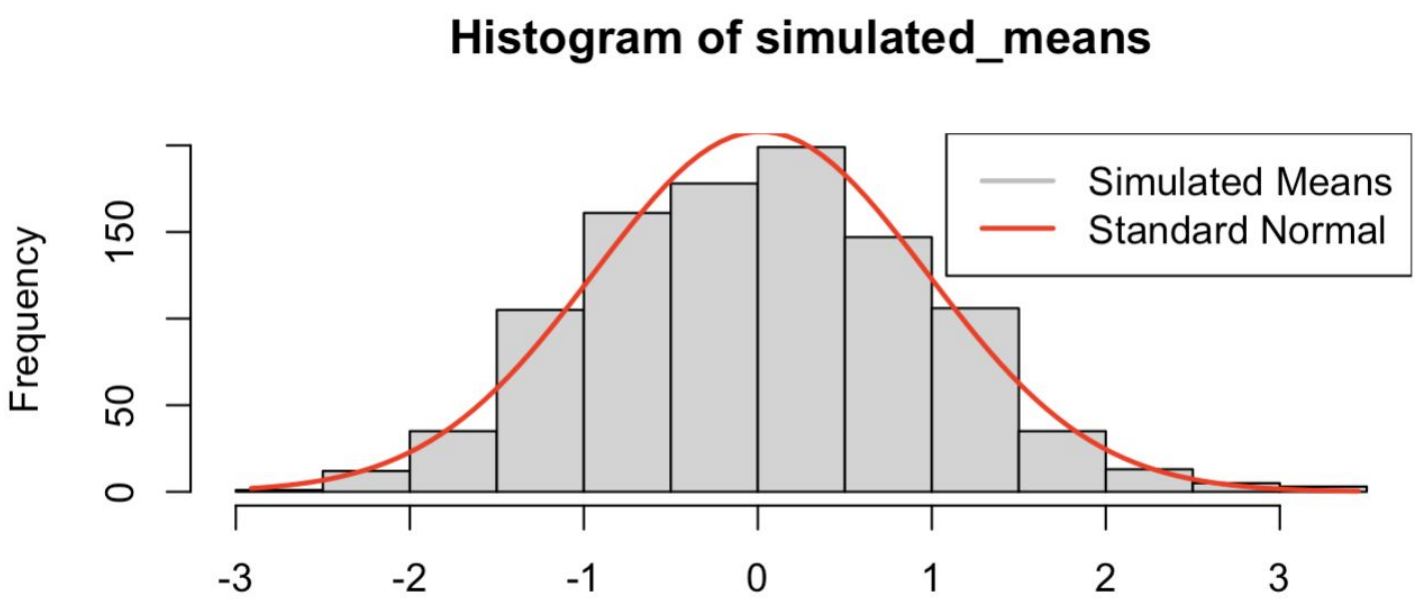


Take a Sample!

Take 10 people from the population and find the mean of their political opinion. But how does this relate to the population mean?

Central Limit Theorem! Bam! Holds regardless of the population distribution.

$$\frac{\sqrt{n}}{\sigma}(\hat{\mu} - \mu) = N(0, 1)$$


Histogram of simulated_means
Frequency
0
50
150
-3
-2
-1
0
1
2
3
Simulated Means
Standard Normal

What if we knew my friend's age?

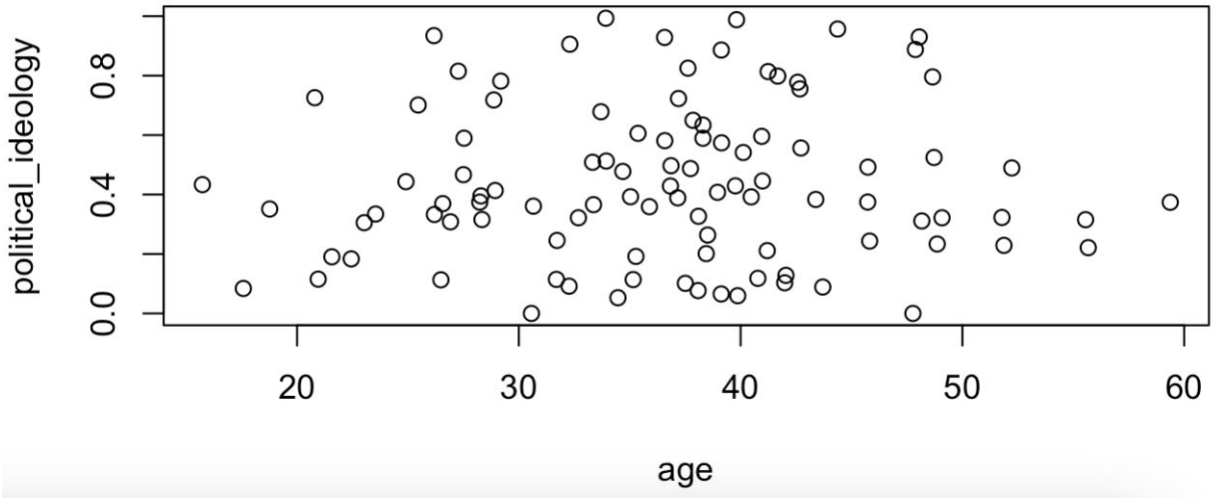
(Without loss of generality for multivariate X)

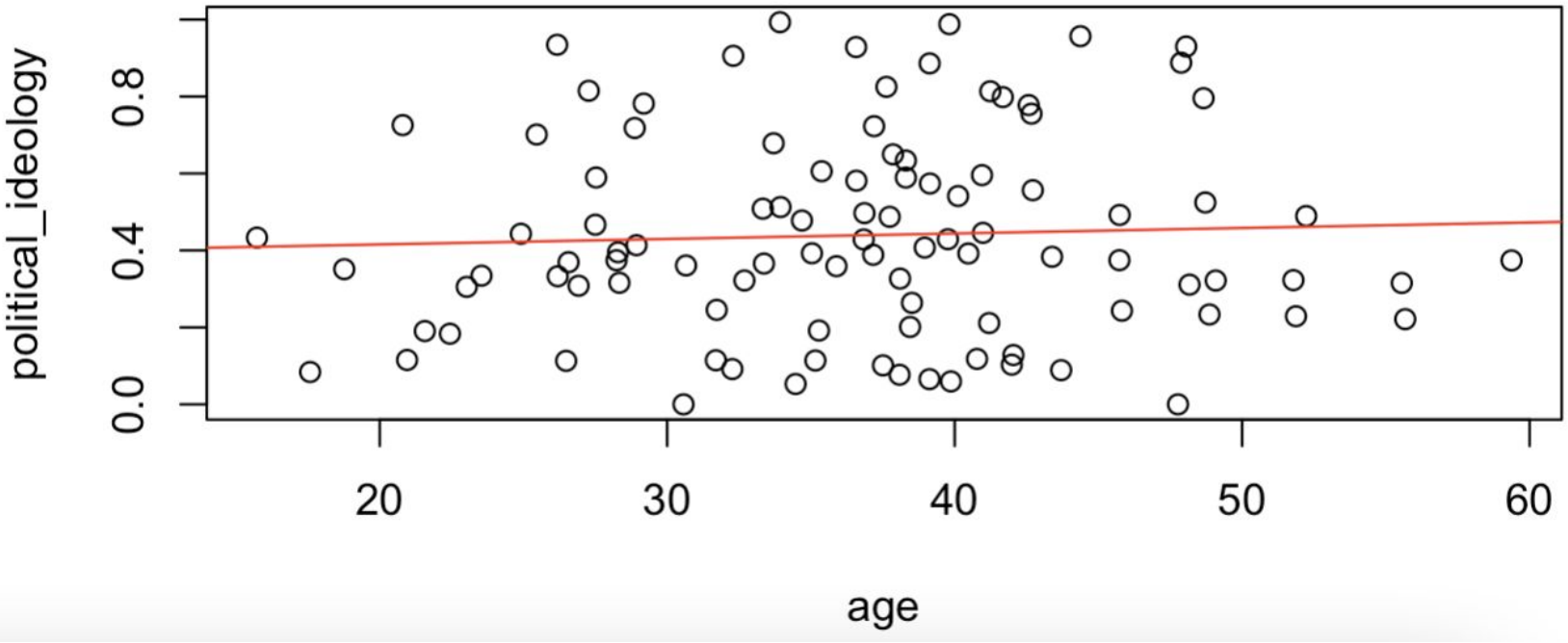
$$\mathbb{E}[Y|X]$$

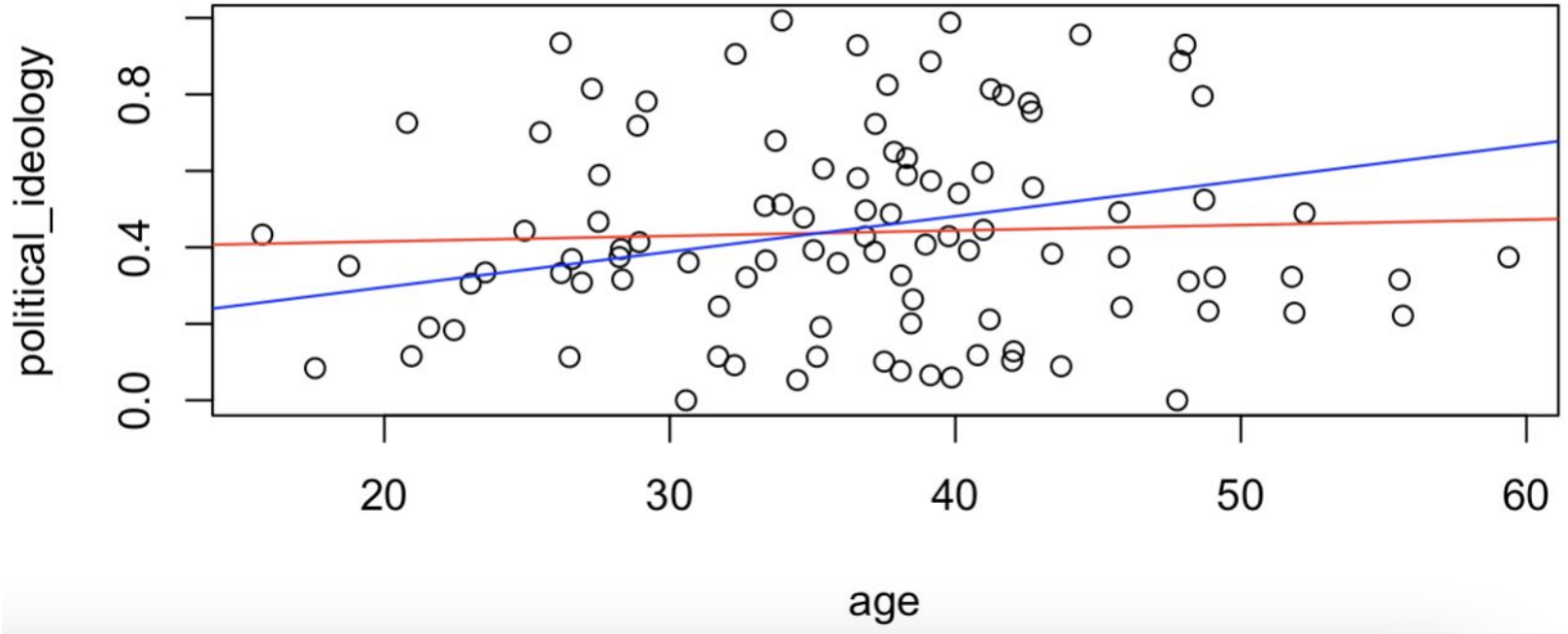
Working model:

$$Y = X\beta + \epsilon$$

Case 1: Political Opinion and Age are Independent

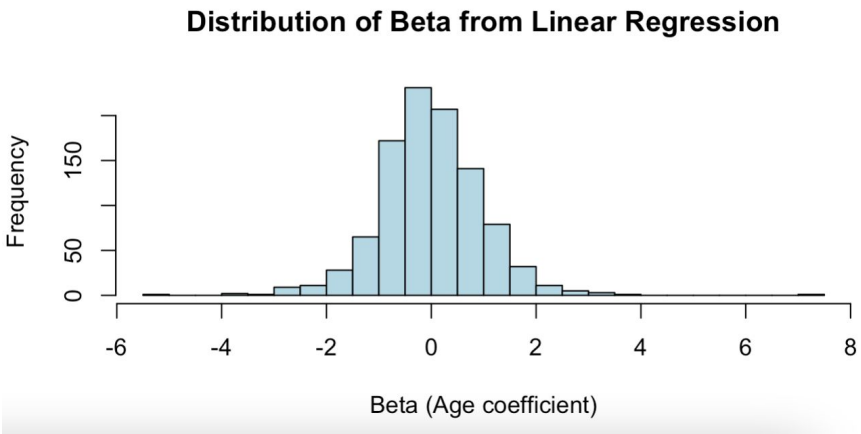


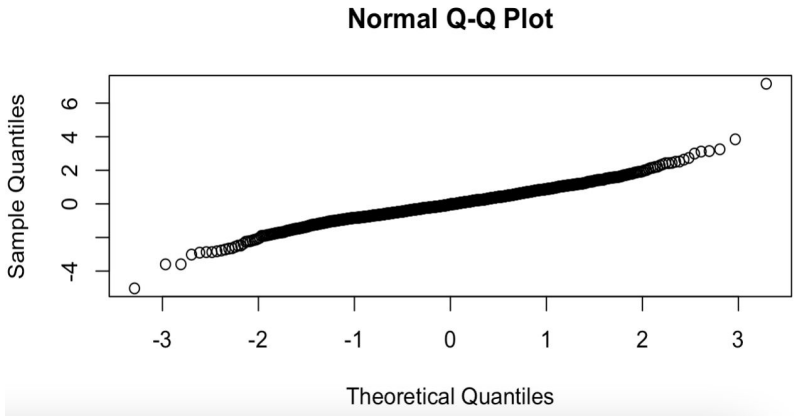

political_ideology
0.0
0.4
0.8
20
30
40
50
60
age


political_ideology
0.0
0.4
0.8
20
30
40
50
60
age

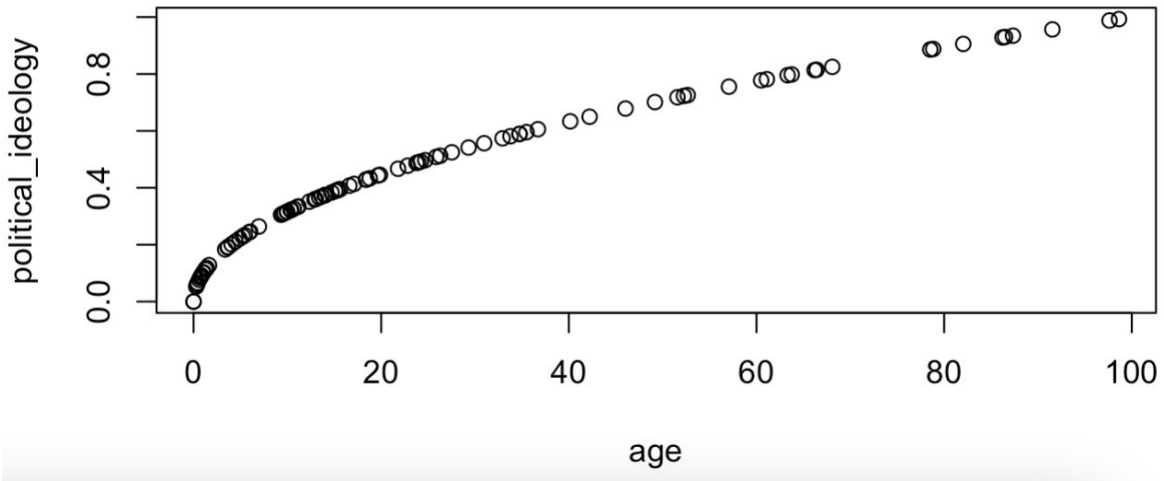
But how does these two lines relate to each other? How good is our prediction?

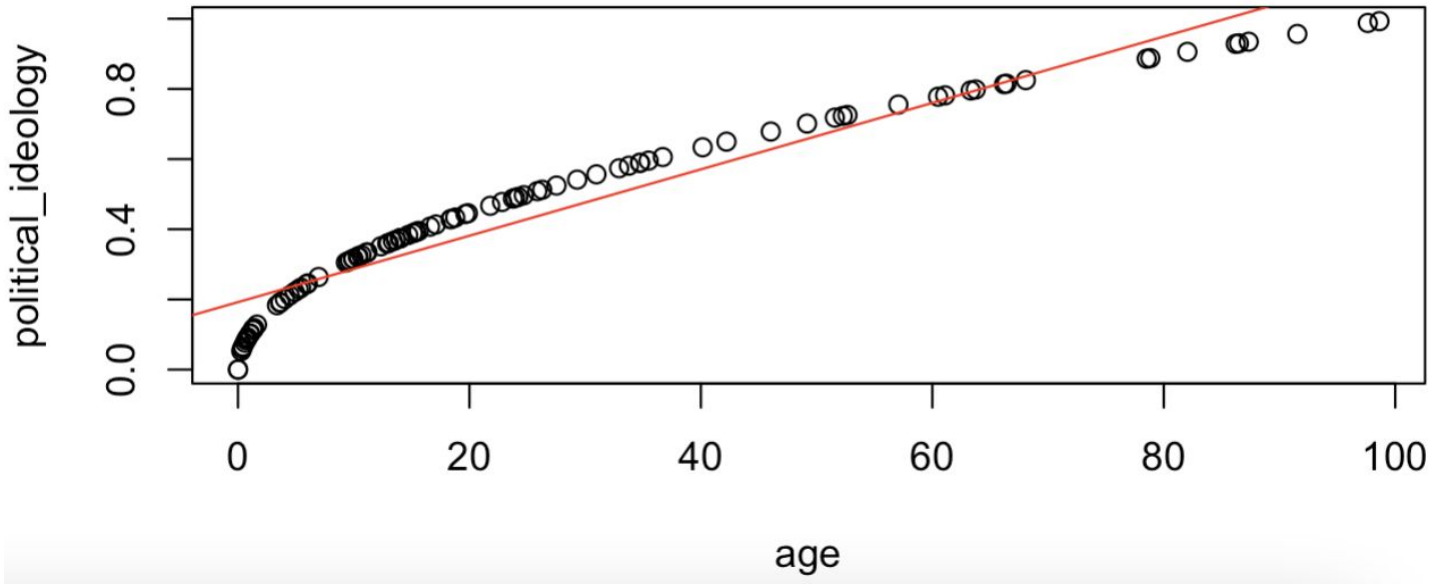
Central Limit Theorem! Bam! Holds regardless of the population distribution.

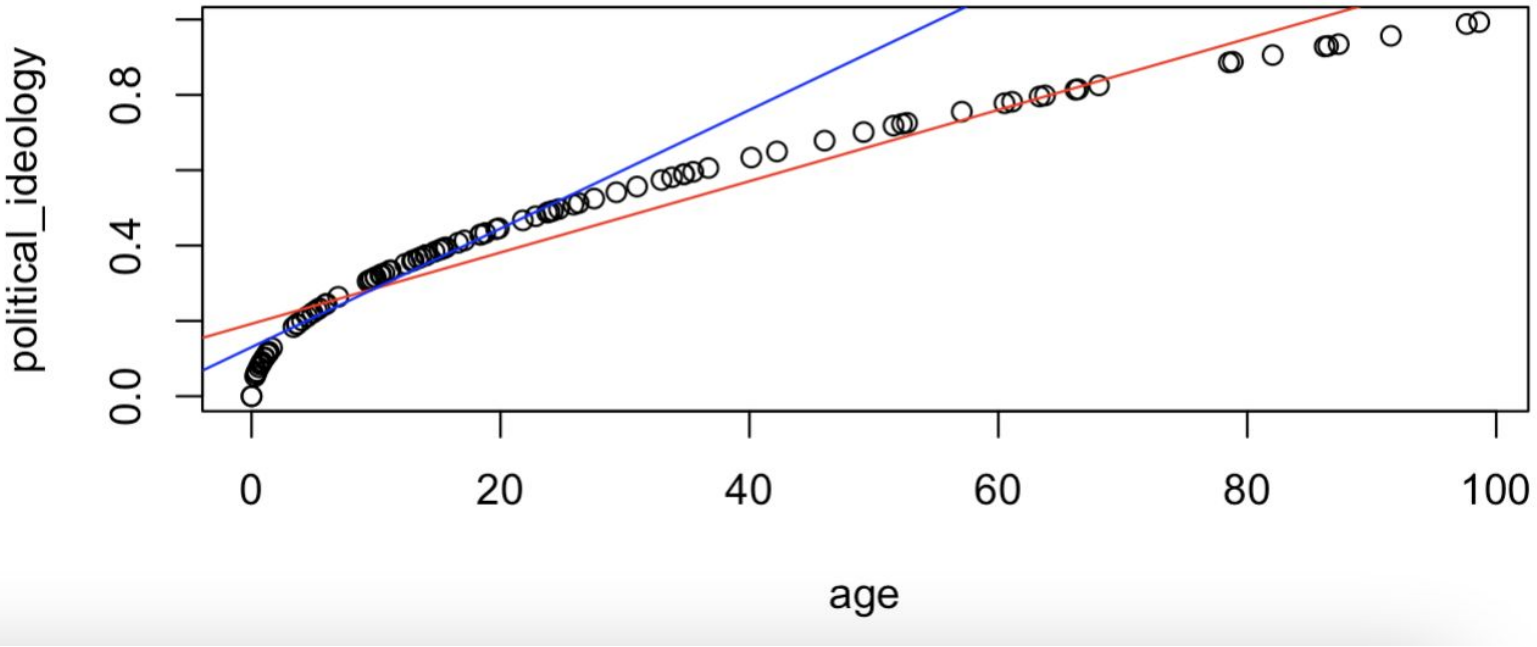




Case 2: Political Opinion and Age are related

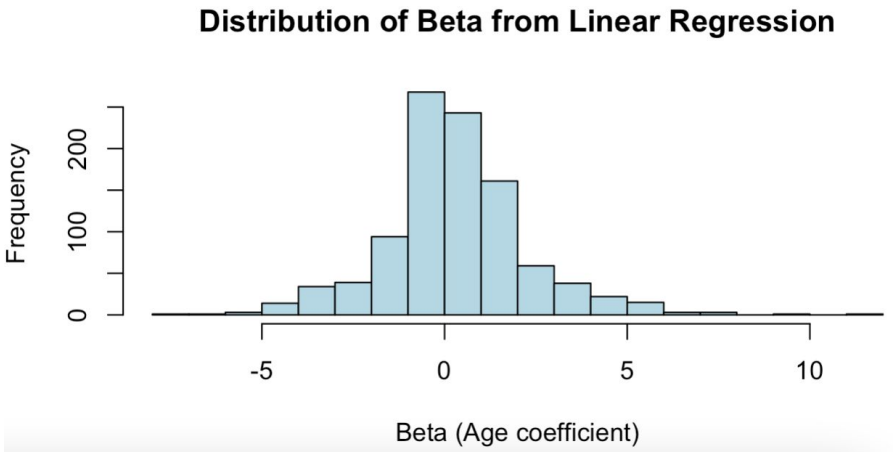


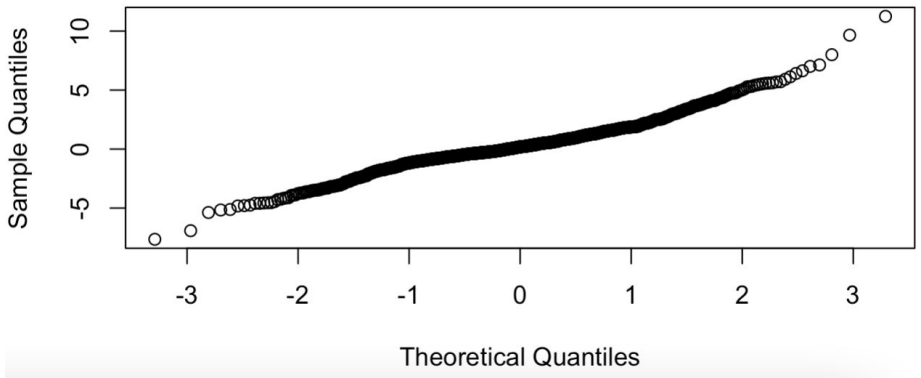

political_ideology
0.0
0.4
0.8
0
20
40
60
80
100
age


political_ideology
0.0
0.4
0.8
0
20
40
60
80
100
age

But how does these two lines relate to each other? How good is our prediction?

Central Limit Theorem! Bam! Holds regardless of the population distribution.





Thank you!