

Dr. Chun's Numb3rs & Løgic

Berkson's Paradox



Young H. Chun, Ph.D.

*Professor of Decision Science &
Cherie H. Flores Endowed Chair of MBA Studies*

Are You Looking for Someone to Love?

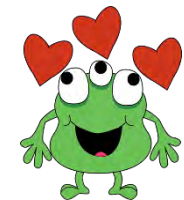


*Top **seven criteria** considered by LSU students when they are looking for...

Rank	Boyfriend	Girlfriend
1	Physical attractiveness	Body
2	Personality	Personality
3	Good general attitude	Money
4	Trustworthiness	Intelligence
5	Wit	Lack of #4
6	Intelligence	Wit
7	Money	Maturity

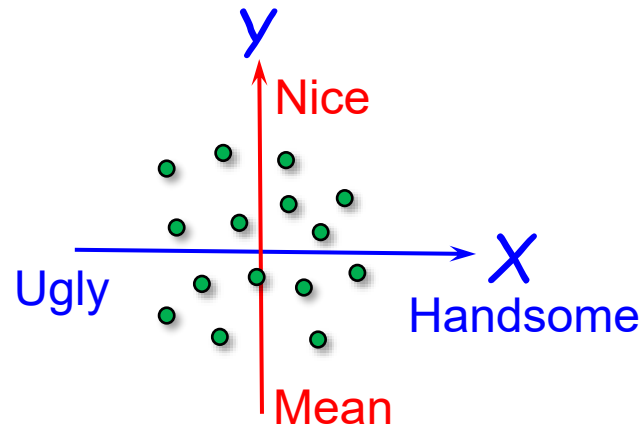
* Reproduced from the *2006 LSU yearbook* without permission

1. Physical attractiveness: Handsome or Ugly
2. Personality: Nice or Mean

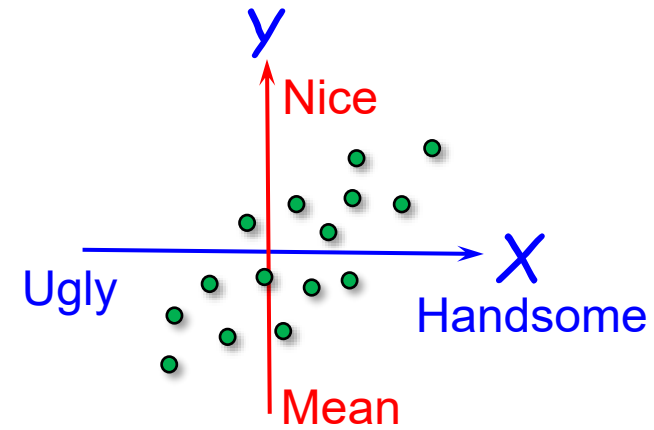




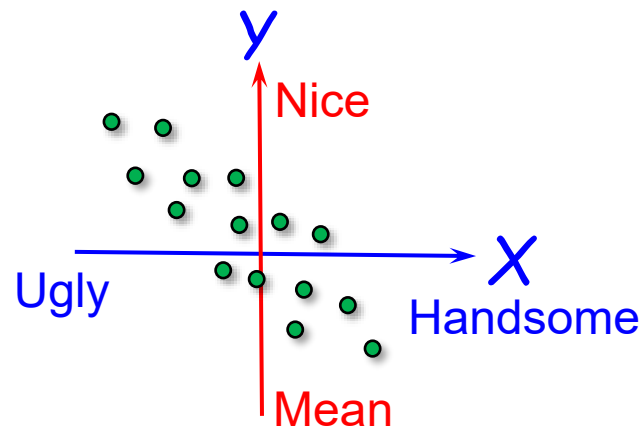
* **Positive** or **Negative** Relationship?



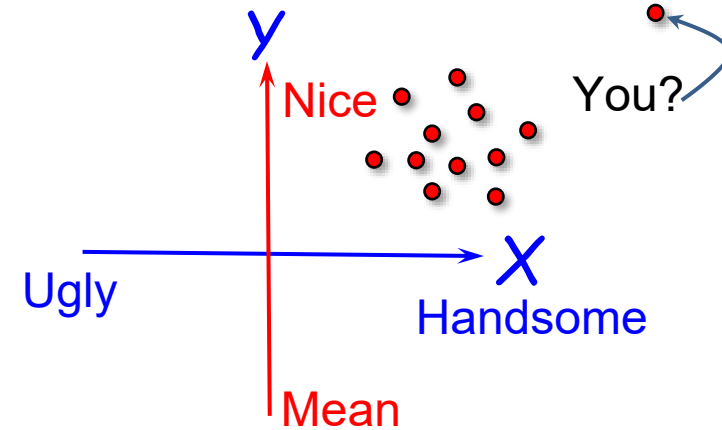
(a) $\text{Corr}[X, Y] = 0$



(b) $\text{Corr}[X, Y] > 0$



(c) $\text{Corr}[X, Y] < 0$



(d) **LSU** MBA Students!

Why Are Handsome Men Such Jerks?

Jordan Ellenberg, *Slate*, June 3, 2014



Suppose you're looking for someone to love. You may have noticed that, among the men in your **dating pool**, the **handsome** ones tend not to be **nice**, and the **nice** ones tend not to be **handsome**.

Is that because having a symmetrical face makes him cruel? Does it mean that being **nice** to people makes him **ugly**? Well, it could be. But it doesn't have to be?

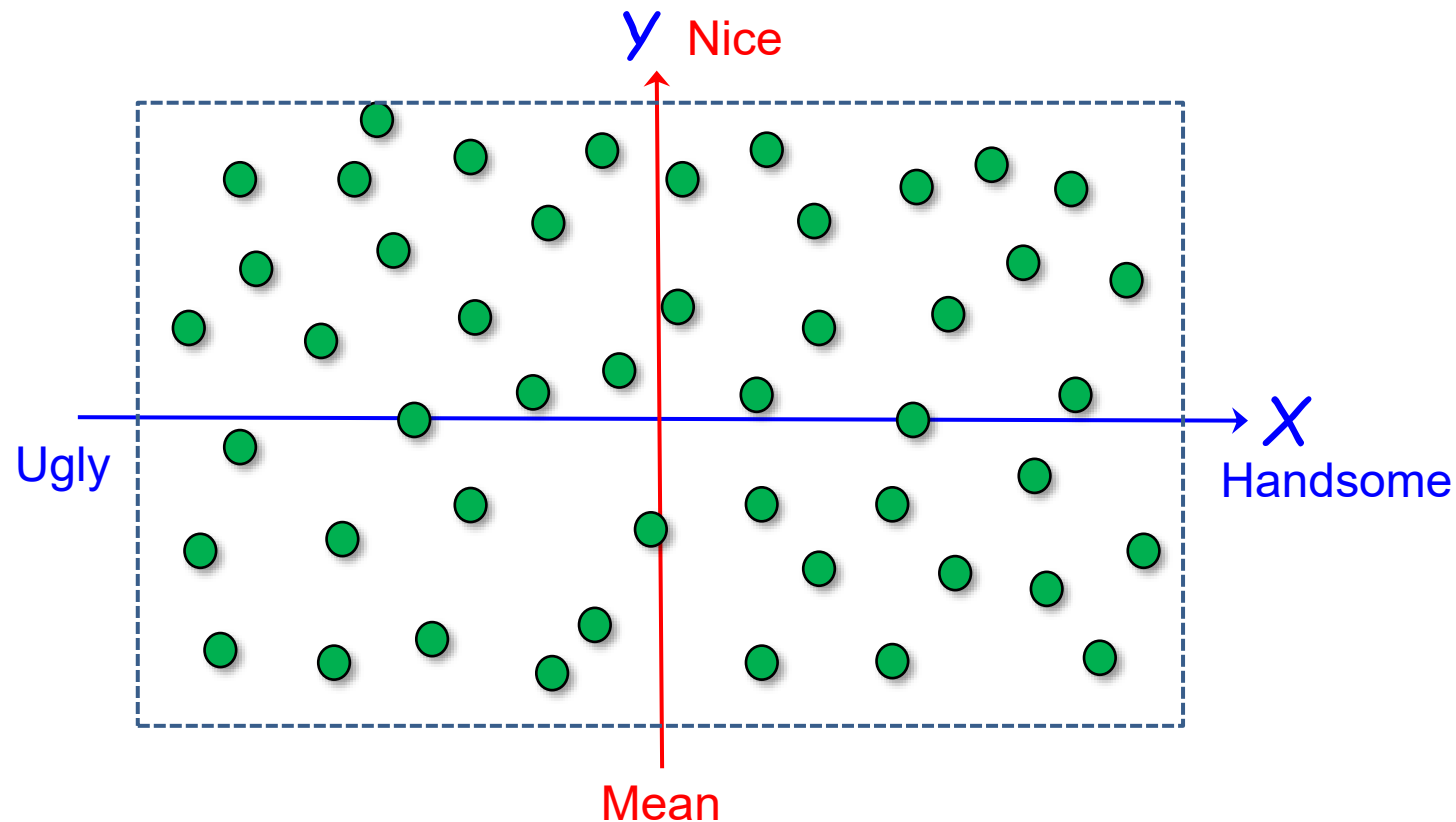
In the dating example, evaluating the relationship between **handsomeness** and **niceness** in your individual optimum **dating pool** gives a false impression of the spurious **negative** relationship between **handsomeness** and **niceness** in the entire potential **dating pool**. Why?





Suppose that men are **equi-distributed** all over this square. In particular, there are nice handsome ones, nice ugly ones, mean handsome ones, and mean ugly ones, in roughly **equal numbers**. (Thus, the **correlation coefficient** between **X** and **Y** is **zero**.)

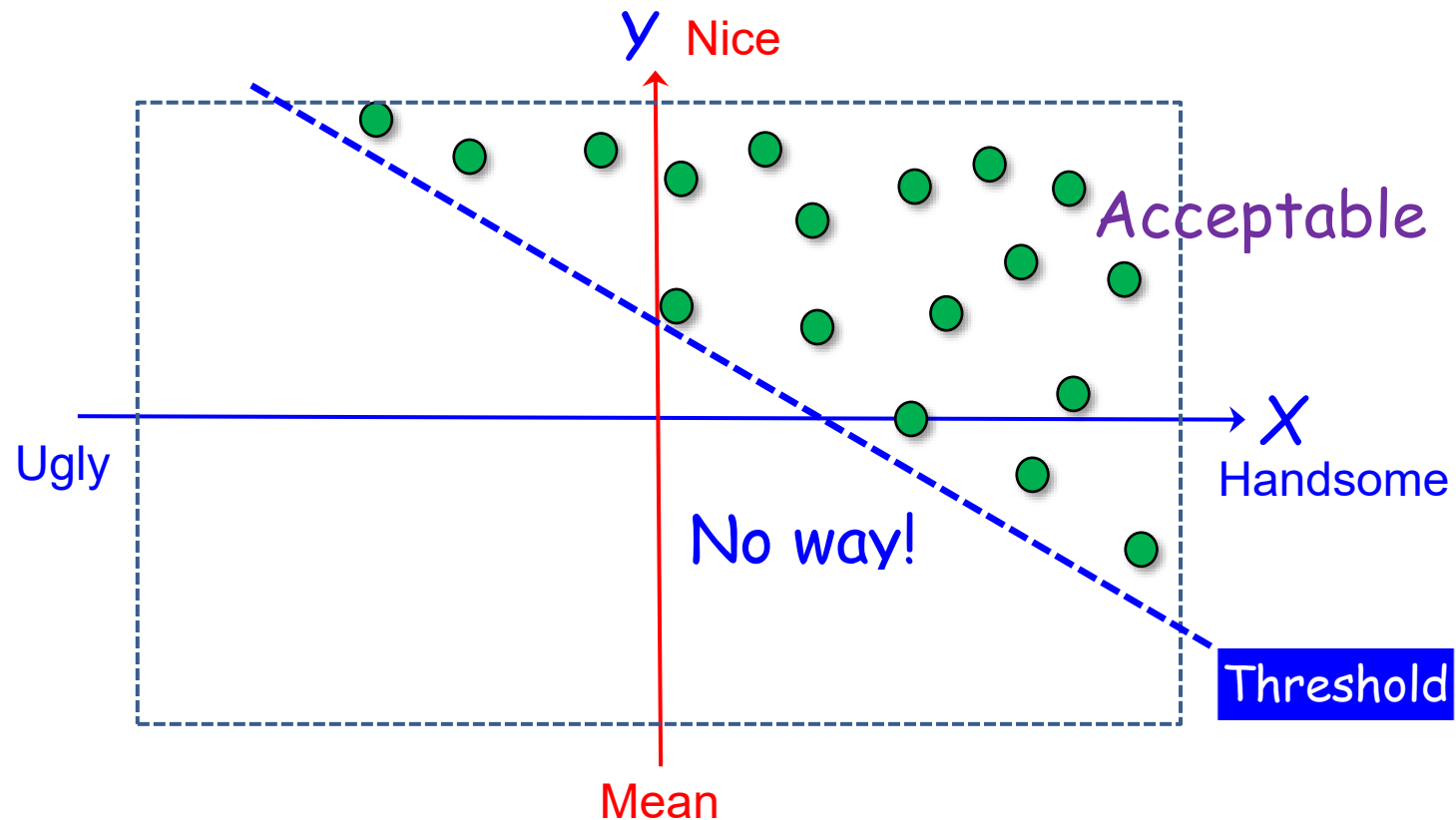
Great Square of Men





Niceness and **handsomeness** have a **common** effect: They put these men in the group of people that **you notice**. Be honest! You will only date a man if his **niceness** plus his **handsomeness** exceeds some **threshold**!

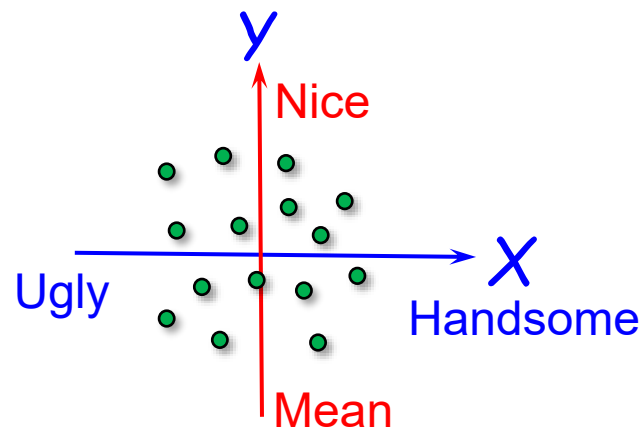
Smaller Triangle of Acceptable Men





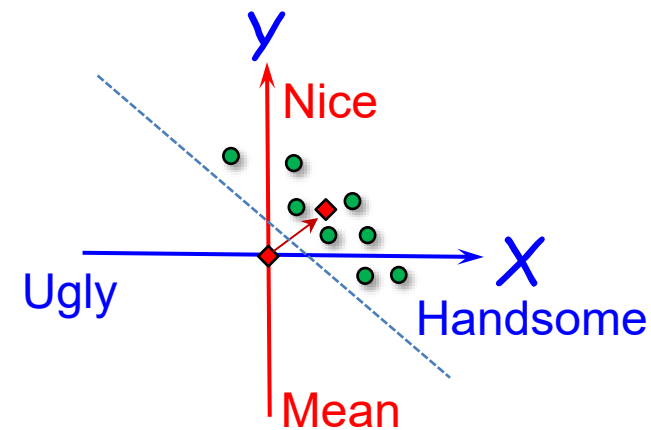
Note that this does not mean that men in your **dating pool** compare unfavorably with men in the **population**. On the contrary, the average nice man that you date is actually more handsome than the average guy in the population.

(a) General population



$$E[X] = 0, E[Y] = 0, \text{Corr}[X, Y] = 0$$

(b) Your dating pool



$$E[X] > 0, E[Y] > 0, \text{Corr}[X, Y] < 0$$

The **negative** correlation between **looks** and **personality** in your **dating pool** is absolutely **real**. But the relation isn't **causal**. If you try to improve your boyfriend's complexion by training him to act **mean**, you've fallen victim to **Berkson's paradox**.



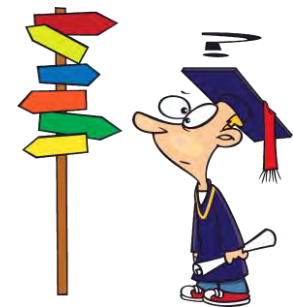
Berkson's Paradox



Berkson's paradox is a particular kind of **selection bias** caused by systematically observing some events more than others. In this paradox, observations are restricted to those where two variables sum together. If you know that $A+B$ must be within a certain range, then having a high A results in a lower B , and vice versa.

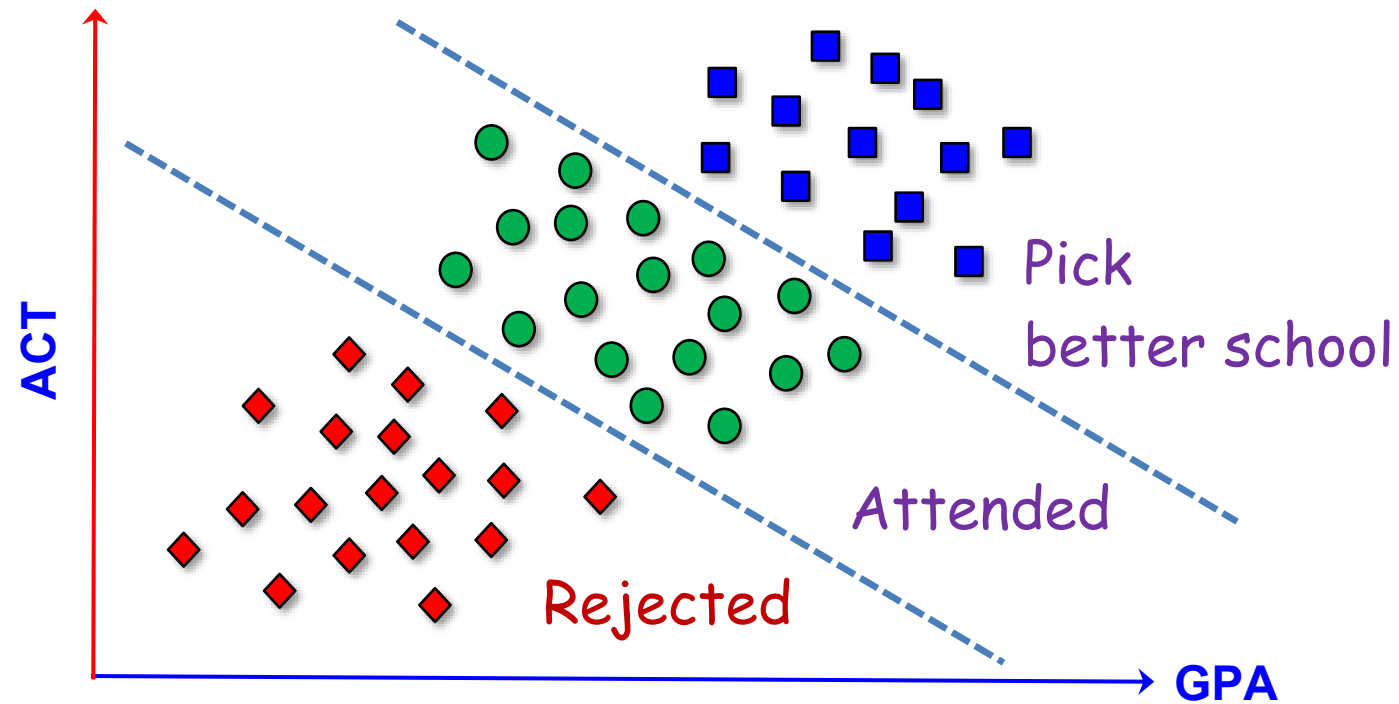
For example, universities pick students based on a number of attributes. In the US, two commonly considered attributes are high school GPA and ACT scores. These are **positively correlated**, so one would expect that **within a given school** they would also be **positively correlated**.

However, this need not be so.





The admissions committee accepts students who have either a sufficiently high **GPA**, a sufficiently high **ACT** score, or some **combination** of the two.



The **range of students** that actually attend the school is given by the green dots in the plot. These dots show a **downward trend**, even though the overall population (red and blue dots) show an **upward trend**. This trend reversal is the "**paradox**," though there is nothing truly paradoxical about it. It is the result of a trade-off between **GPA** and **ACT** scores in the applicants reviewed.

Berkson's Restaurants

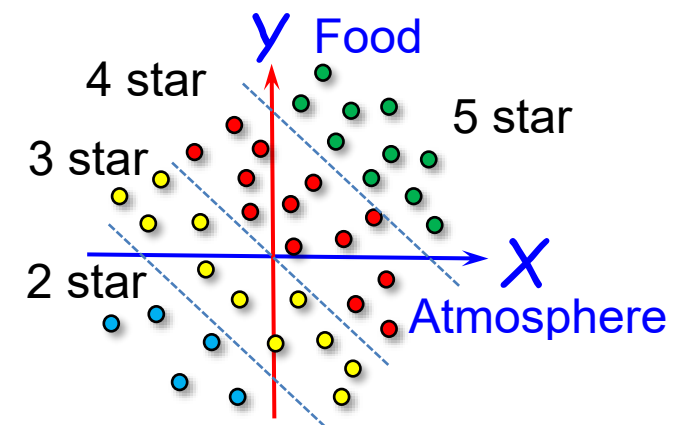


Patrons rate **restaurants** based on a number of features, including (i) the quality of the **food** and (ii) the quality of the "**atmosphere**" (what it looks like, what music they play, etc). They combine these ratings into an **overall rating** from **one** to **five** stars.

You come across two restaurants that you know have **five star ratings**. You can see inside, but you haven't tasted the food yet. Which restaurant is more likely to have the **best tasting food**?



- (a) Equally likely
- (b) Worse-looking restaurant
- (c) Better-looking restaurant



Movie Trivia



A **rat** who can cook makes an unusual alliance with a young **kitchen worker** at a famous **restaurant**.

Ratatouille (2007)



A **rat** who can cook makes an unusual alliance with a young **kitchen worker** at a famous **restaurant**.