



發掘有趣的問題 -- 資料科學案例介紹

何宗武

臺師大管理學院 全營所

tsungwu@ntnu.edu.tw

<http://web.ntnu.edu.tw/~tsungwu/>



國立臺灣師範大學
NATIONAL TAIWAN NORMAL UNIVERSITY

請問資料科學，要如何確認這些命題？

「善有善報，惡有惡報」
「官大學問大」

資料科學方法論的初步回答

1. 要怎麼確認？Re-formulate the problem
2. 須要何種資料？Measure your idea

資料有兩種

- Survey
- Non-Survey



國立臺灣師範大學
NATIONAL TAIWAN NORMAL UNIVERSITY

MLB 精彩賽事

which seasons were unlike any others ?

https://www.kdnuggets.com/2019/09/time-series-baseball.html?fbclid=IwAR3iqkugPpM8u7kEIFg_43OvPIPQt3P6yRkjh7-XO4VH5GpfORJJfrGOV7k



國立臺灣師範大學
NATIONAL TAIWAN NORMAL UNIVERSITY

What are you going to do?

- Are Books Better Than the Movies?
- Are Black Americans Killed more than People in other Developed Countries?
- Can Smiling Really Make You Happier?
- Which Movies Make Grown Men Cry?

資料科學方法論的初步回答

1. 要怎麼確認？
2. 須要何種資料？

資料有兩種

- Survey
- Non-Survey



國立臺灣師範大學
NATIONAL TAIWAN NORMAL UNIVERSITY

說的比唱的好聽？



關於小說和電影，可以提出什麼問題？

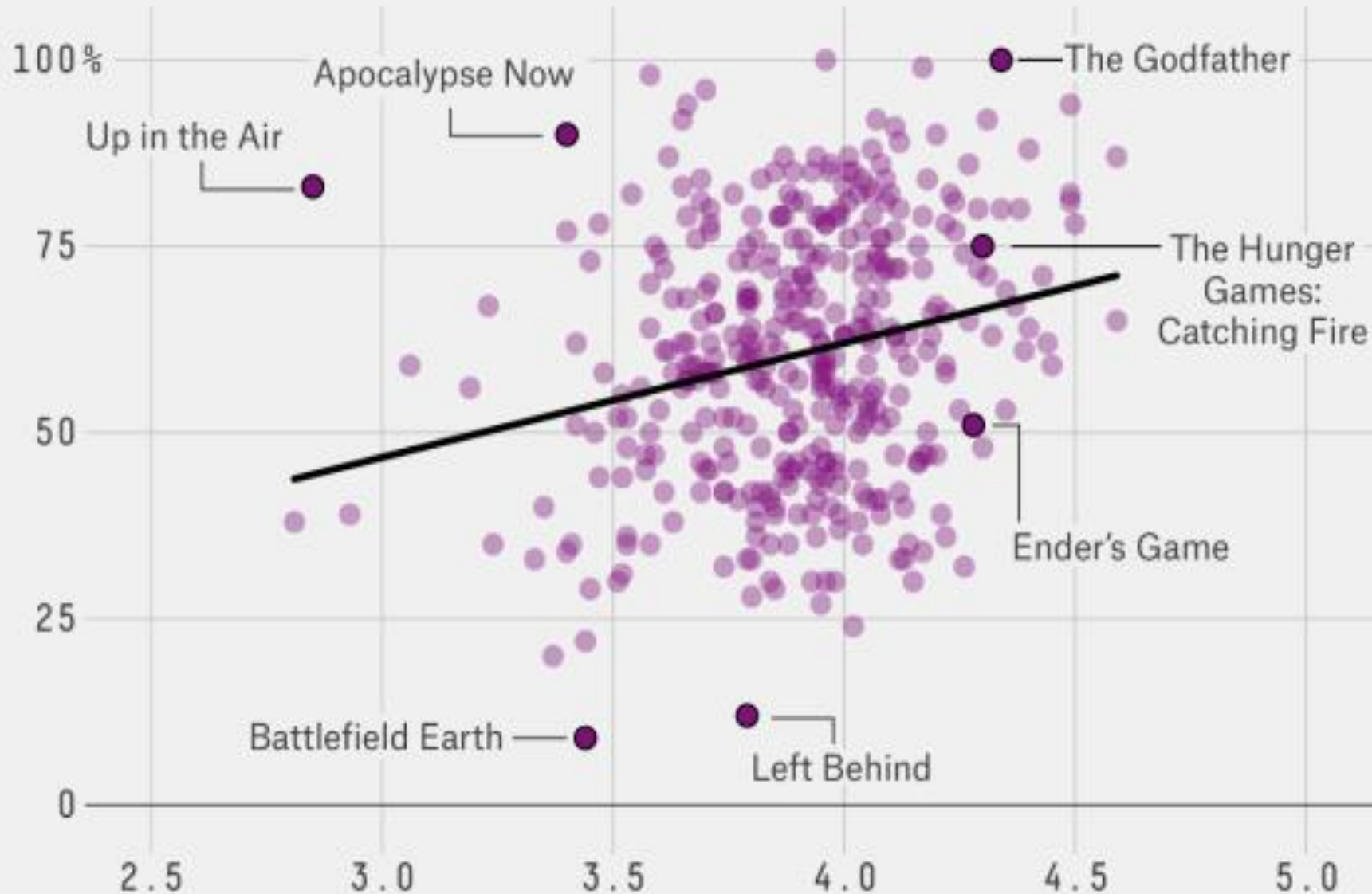
1. 哪些電影拍的比原著好？
2. 哪些原著因為電影翻紅？



國立臺灣師範大學
NATIONAL TAIWAN NORMAL UNIVERSITY

When Books Become Movies

Metacritic score of films vs. Goodreads score of source novel



國立臺灣師範大學
NATIONAL TAIWAN NORMAL UNIVERSITY

RANK	FILM	METACRITIC SCORE	GOODREADS RATING	Z-SCORE DIFF.
1	Up in the Air (2009)	83	2.9	5.3
2	Apocalypse Now (1979)	90	3.4	3.6
3	Metropolis (1927)	98	3.6	3.4
4	Scent of a Woman (1993)	59	3.0	3.1
5	Mr. Holmes (2015)	67	3.2	2.9
6	Sideways (2005)	94	3.7	2.8
7	The Graduate (1967)	77	3.4	2.8
8	Taking Lives (2004)	38	2.8	2.8
9	Dr. Strangelove (1964)	96	3.7	2.8
10	There Will Be Blood (2008)	92	3.7	2.8



FILM	METACRITIC SCORE	GOODREADS RATING	Z- SCORE DIFF.
1 Addicted (2014)	32	4.3	-3.0
2 Vampire Academy (2014)	30	4.2	-2.7
3 The Jackal (1997)	36	4.2	-2.6
4 Safe Haven (2013)	34	4.2	-2.5
5 The Last Song (2010)	33	4.1	-2.4
6 The Mortal Instruments: City of Bones (2013)	33	4.1	-2.4
7 Percy Jackson: Sea of Monsters (2013)	39	4.2	-2.4
8 Left Behind (2014)	12	3.8	-2.4
9 A Walk to Remember (2002)	35	4.1	-2.4
10 Harry Potter and the Deathly Hallows: Part 1 (2010)	65	4.6	-2.3
11 Divergent (2014)	48	4.3	-2.2
12 Unbroken (2014)	59	4.5	-2.1
13 The Book Thief (2013)	53	4.4	-2.1
14 Mortdecai (2015)	27	4.0	-2.1
15 Where the Heart Is (2000)	30	4.0	-2.1
16 Pay It Forward (2000)	40	4.1	-2.0
17 The Three Musketeers (2011)	35	4.0	-2.0
18 Seventh Son (2015)	30	4.0	-2.0
19 Ender's Game (2013)	51	4.3	-2.0
20 The Help (2011)	62	4.4	-1.9

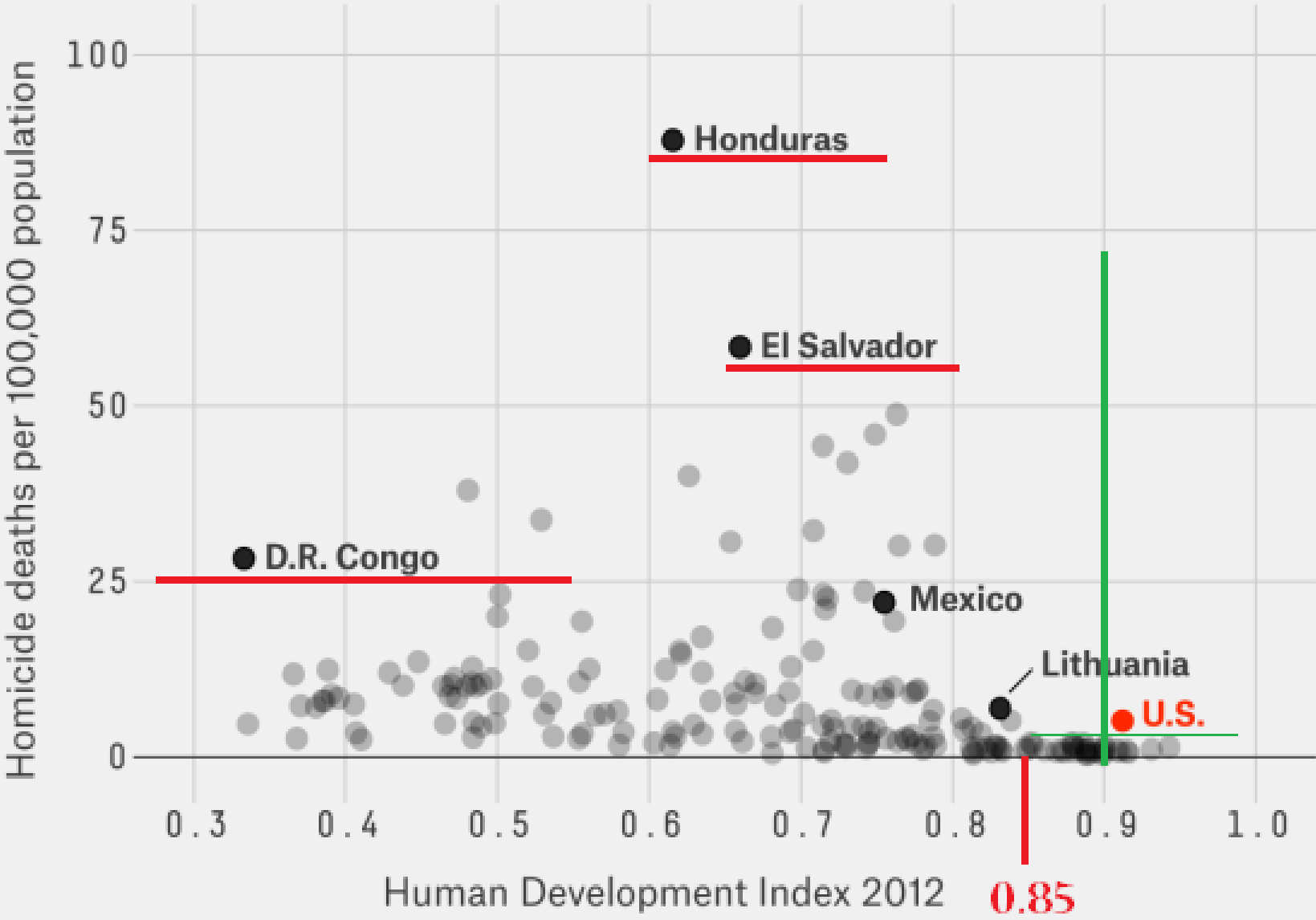
Can you propose similar question?



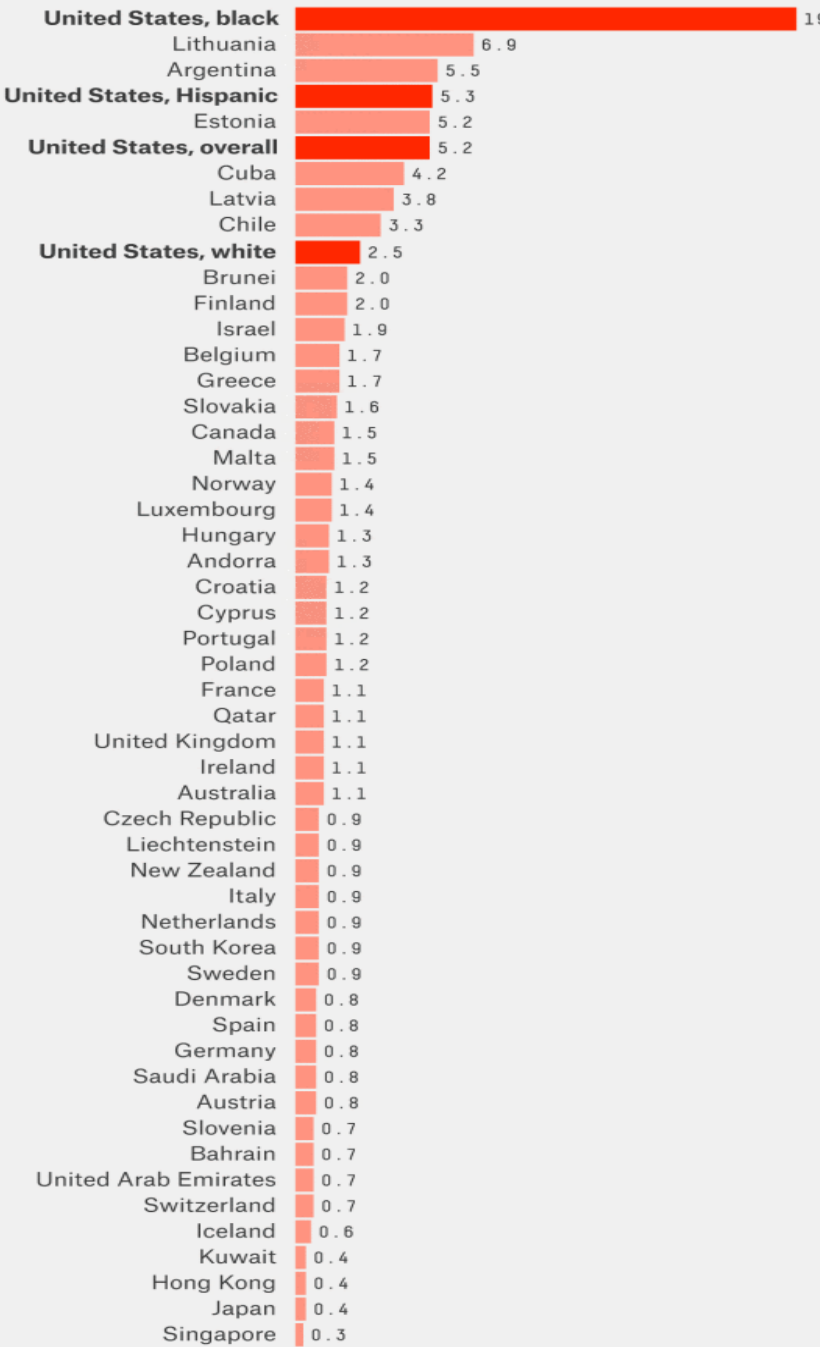
國立臺灣師範大學
NATIONAL TAIWAN NORMAL UNIVERSITY

Is The U.S. An Outlier In Homicides?

Homicide deaths per 100,000 population, 2010-12 average

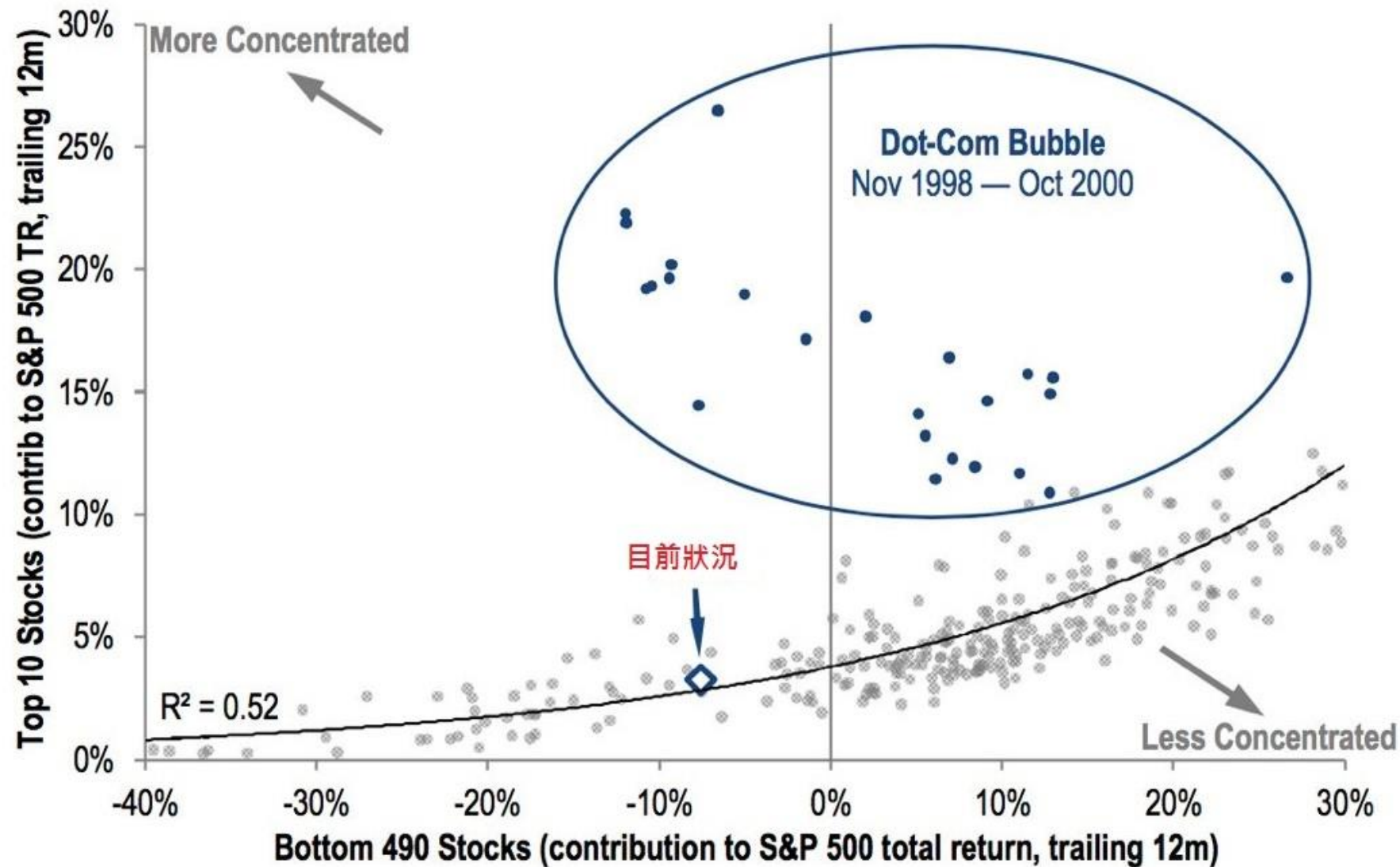


Homicide Deaths Per 100,000 Population
2010-12 average, high-Human Development Index countries



Total Return Contribution: Top 10 vs Bottom 490 Stocks

12m rolling total return, monthly since 1990



猜猜看：這張圖在說什麼？

Source: J.P. Morgan, S&P and Bloomberg



國立臺灣師範大學
NATIONAL TAIWAN NORMAL UNIVERSITY

視覺化的案例

民主黨總統候選人第三次辯論

<https://fivethirtyeight.com/features/the-third-democratic-debate-in-7-charts/>



國立臺灣師範大學
NATIONAL TAIWAN NORMAL UNIVERSITY

開卷有益

請問資料科學，要如何確認這個因果命題？

「善有善報，惡有惡報」

「不是不報，時候未到」



國立臺灣師範大學
NATIONAL TAIWAN NORMAL UNIVERSITY