

Chicken soap

HUST Bioinformatics course series for undergraduates

Wei-Hua Chen (CC BY-NC 4.0)

01 September, 2023

Section 1 : outline

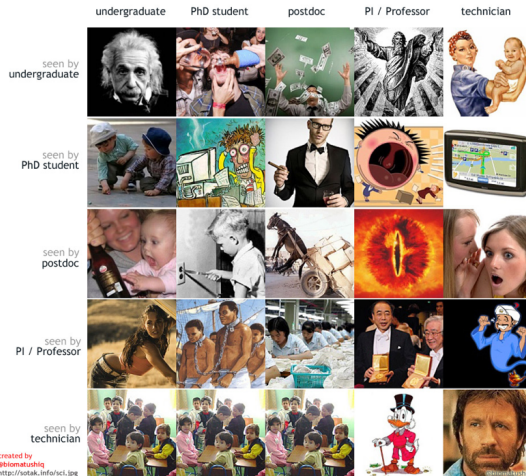
Outline

- the two most important figures in science
- think in English
- think creative
- find your own truth

section 2 : contents

how people in science see each other

How people in science see each other



the most important Scientific virtue



Figure 2: most important core values selected by elite scientists

think in English

怎么练习口语和语感？

- 1 Find a short article that you know every word
- 2 Read it aloud at a quiet place so that you can hear it
- 3 Do this everyday for **three years**

think creative

How?

- 1 Read a research paper
- 2 Find three limitations of the research
- 3 Address these issues with your own ideas
- 4 Do this once a week for **three years**

try find your own truth

未经调研，不要轻易相信别人结论，特别是科普文章。

There is a reproducibility crisis in scientific community:

- 70% of researchers failed to reproduce others' experiments
- 50% failed to reproduce their own

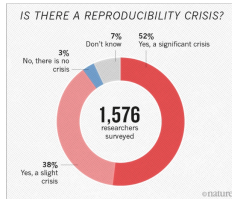
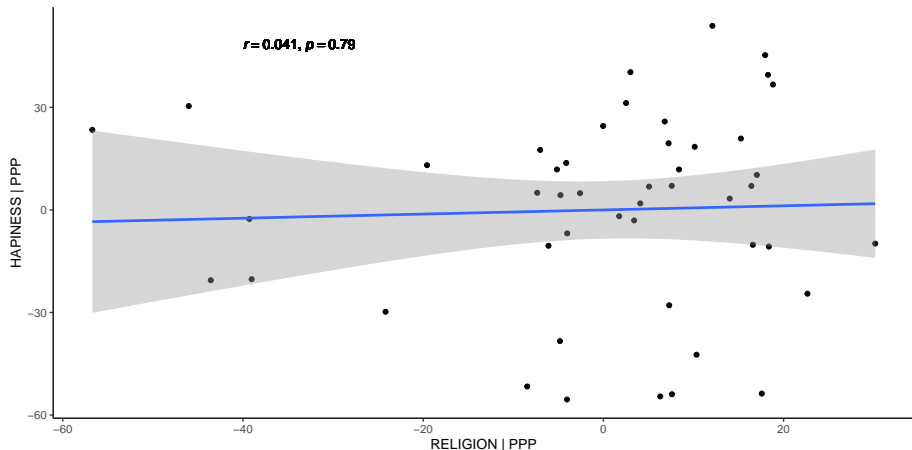


Figure 3: reproducibility crisis (Nature 2015 survey)

Religious and happiness

Are religious people happier? (Pew research 2019)

Religion, GDP and happiness.



The data

```
data/talk00/data.xlsx
```

```
head(rhp);
```

```
## # A tibble: 6 x 4
##   country    rel happiness    ppp
##   <chr>    <dbl>      <dbl> <dbl>
## 1 Ghana      96        207.  4266
## 2 Nigeria    93        183.  6108
## 3 Armenia    92        123.  8468
## 4 Fiji       92        223.  9044
## 5 Romania    89        173. 20787
## 6 Kenya    88        187.  3208
```

Single factor analysis

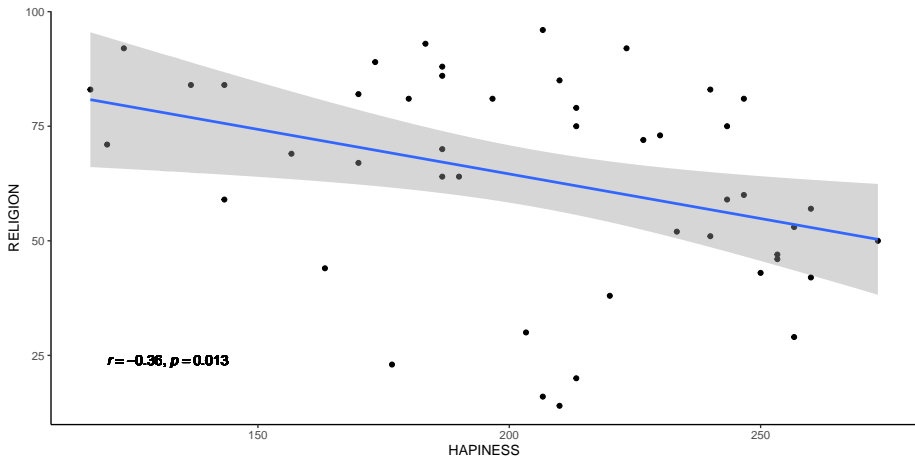
Happiness vs. Religion

```
c = cor.test( rhp$happiness, rhp$rel);
eq <- substitute(
  italic(r)~"="~r2*","~italic(p)==pvalue,
  list(
    r2 = as.vector( format( c$estimate , digits = 2) ),
    pvalue = as.vector( format( c$p.value , digits = 2) )
  ));
eq <- as.character(as.expression(eq));

m<-ggplot(rhp, aes(x=happiness, y=rel)) +
  geom_point() +
  labs(y="RELIGION", x = "HAPINESS")+
  theme_classic() +
  geom_smooth(method=lm) +
  geom_text( data = NULL,
    aes( x = 120, y = 25, label= eq, hjust = 0, vjust = 1),
    size = 4, parse = TRUE, inherit.aes=FALSE);
```

Plot

m



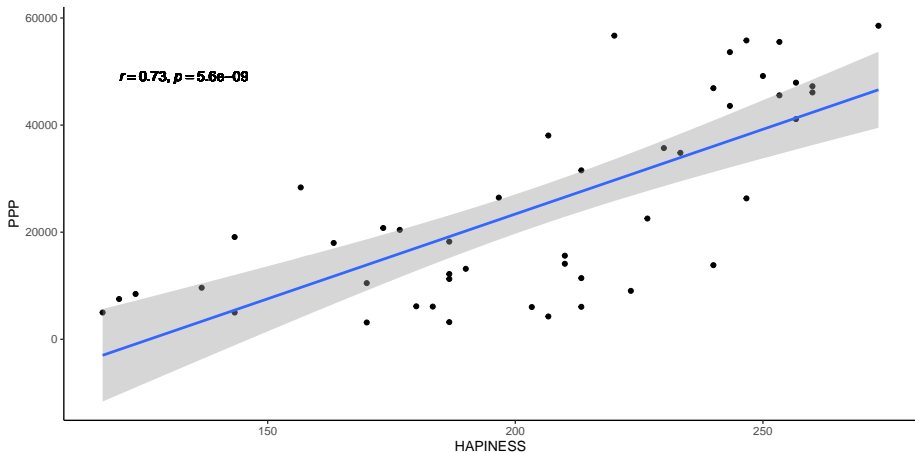
Happiness vs. ppp

```
c = cor.test( rhp$happiness, rhp$ppp);
eq <- substitute(
  italic(r) ~ "=" ~ r2 * ", " ~ italic(p) == pvalue,
  list(
    r2 = as.vector( format( c$estimate , digits = 2) ),
    pvalue = as.vector( format( c$p.value , digits = 2) )
  ));
eq <- as.character(as.expression(eq));

m2 <- ggplot(rhp, aes(x=happiness, y=ppp)) +
  geom_point() +
  labs(y="PPP", x = "HAPPINESS") +
  theme_classic() +
  geom_smooth(method=lm) +
  geom_text( data = NULL,
    aes( x = 120, y = 50000, label= eq, hjust = 0, vjust = 1),
    size = 4, parse = TRUE, inherit.aes=FALSE);
```

Plot

m2



Religion vs. ppp

```

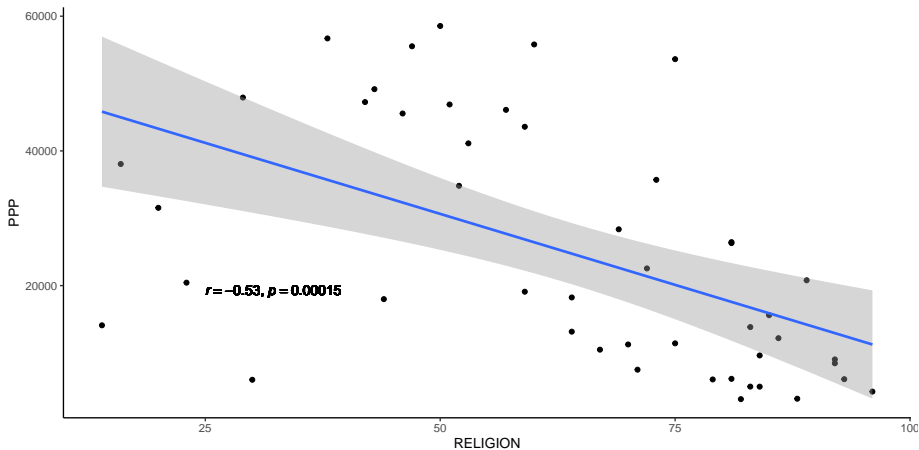
c = cor.test( rhp$rel, rhp$ppp);
eq <- substitute(
  italic(r) ~ "=" ~ r2 * ", " ~ italic(p) == pvalue,
  list(
    r2 = as.vector( format( c$estimate , digits = 2) ),
    pvalue = as.vector( format( c$p.value , digits = 2) )
  ));
eq <- as.character(as.expression(eq));

m3 <- ggplot(rhp, aes(x=rel, y=ppp)) +
  geom_point() +
  labs(y="PPP", x = "RELIGION") +
  theme_classic() +
  geom_smooth(method=lm) +
  geom_text( data = NULL,
    aes( x = 25, y = 20000, label= eq, hjust = 0, vjust = 1),
    size = 4, parse = TRUE, inherit.aes=FALSE);

```


Plot

m3



Happiness vs. PPP vs. Religion

```
summary( glm( happiness ~ ppp + rel, data = rhp ) );
```

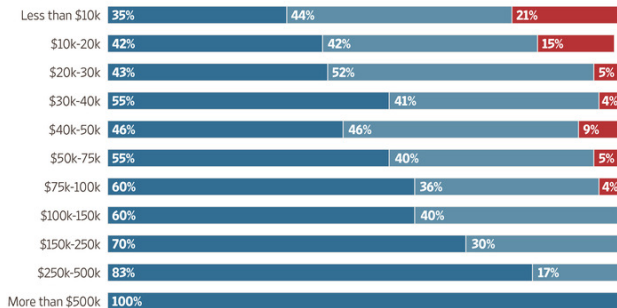
```
##
## Call:
## glm(formula = happiness ~ ppp + rel, data = rhp)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -55.220  -14.868    5.049   18.523   53.136
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.580e+02  1.932e+01  8.175 2.26e-10 ***
## ppp          1.725e-03  2.791e-04  6.182 1.82e-07 ***
## rel          6.048e-02  2.239e-01  0.270  0.788
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 840.0595)
##
##      Null deviance: 79371  on 46  degrees of freedom
## Residual deviance: 36963  on 44  degrees of freedom
## AIC: 454.75
##
## Number of Fisher Scoring iterations: 2
```

How happy are you???

How Happy Are You?

In a study by Betsey Stevenson and Justin Wolfers at the University of Michigan, the relationship between happiness and income in the U.S. was analyzed.

Happiness ■ Very happy ■ Fairly happy ■ Not too happy



Source: "Subjective Well-Being and Income: Is There Any Evidence of Satiation?" paper by Betsey Stevenson and Justin Wolfers, University of Michigan

Note: Some categories may not equal 100% due to rounding.

The Wall Street Journal

Figure 4: Data source: WSJ "Can Money Buy You Happiness?"

Concluding remarks

- 不轻信、重事实
- 有创新性
- 有用