



PSYC122 week 11: Correlation



Dr. Margriet A. Groen

Outline

- What is correlation? Correlation coefficient and scatterplots
- Types of correlation Interpreting correlation coefficient and scatterplots
- Covariance and correlation How correlation is derived from covariance
- Hypothesis testing Critical values for significance
- Coefficient of determination R^2
- Why correlation does not infer causation
- How to conduct correlation **R**
- How to report correlation **APA**

The data

Case	Abil	IQ	Home	TV
1	61	107	144	487
2	56	109	123	608
3	45	81	108	640
4	66	100	155	493
5	49	92	103	636
6	62	105	161	407
7	61	92	138	463
8	55	101	119	717
9	62	118	155	643
10	61	99	121	674
11	51	104	93	675
12	48	100	127	595
13	50	95	97	673
14	50	82	140	523
15	67	114	151	665
16	51	95	112	663
17	55	94	102	684
18	54	103	142	505
19	57	96	127	541

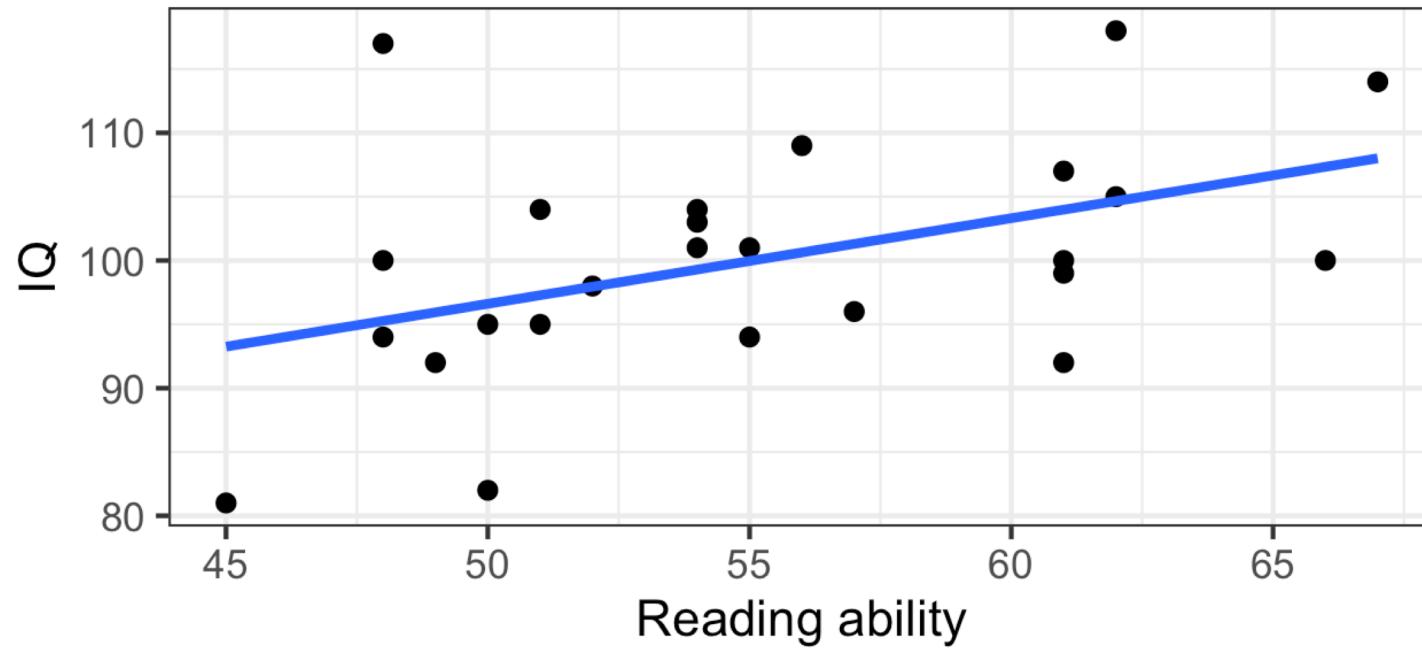
```
library("broom")
library("tidyverse")

# read in the data
mh <- read_csv("MillerHadenData.csv")
mh
```

How to make a scatterplot

```
# plot the relationship between reading ability and IQ using a scatterplot and a line of best fit
ggplot(mh, aes(x = Abil, y = IQ)) +
  geom_point() +
  geom_smooth(method = "lm", se = FALSE) +
  theme_bw() +
  labs(x = "Reading ability")
```

Scatterplot output



How to conduct correlation analysis in R

```
# conduct a correlation analysis, using Pearson's r
results <- cor.test(mh$Abil,
                     mh$IQ,
                     method = "pearson",
                     alternative = "two.sided") %>%
tidy()

results
```

Output in R

```
> results
# A tibble: 1 × 8
  estimate statistic p.value parameter conf.low conf.high method      alternative
    <dbl>     <dbl>    <dbl>      <int>     <dbl>     <dbl> <chr>
1 0.451      2.42  0.0236       23     0.0682     0.718 Pearson's product-moment correlation two.sided
```

estimate Pearson's *r*
p.value p-value
parameter degrees of freedom

Reporting correlation analysis (*cont.*)

What you need to include:

- Correlation coefficient
- Degrees of freedom (N-2)
- Significance value

```
# pull out Pearson's r, the degrees of freedom and the p-value for reporting the results
r <- results %>%
  pull(estimate) %>%
  round(2)

df <- results %>%
  pull(parameter)

pvalue <- results %>%
  pull(p.value) %>%
  round(3)
```

Reporting correlation analysis

Example: A Pearson's correlation coefficient was used to assess the relationship between reading ability and intelligence. There was a significant positive correlation, $r(23) = .45, p = .024$. As reading ability increased, intelligence increased.

Example of a correlation matrix (APA format)

Table 1. *A correlation matrix showing the relationship between reading ability and motivation.*

	Reading ability	Intelligence
Reading ability	-	-
Intelligence	.45*	-

* $p < .05$