

Exploring Dr. Lynn’s 2015 Dataset on Race Differences in Intelligence

Julia Kim

```
#### Preamble ####
# Purpose: Read in Dr. Lynn's 2015 dataset on race differences in
  ↳ intelligence # and make a graph. Relate this dataset to the quote by
  ↳ Dr. Srinivasan.
# Author: Julia Kim
# Email: juliaym.kim@mail.utoronto.ca
# Date: 18/02/24
# Prerequisites: Know where to get the 2015 dataset on race differences
  ↳ in intelligence, be familiar with the quote by Dr. Srinivasan.
```

The question of race differences in intelligence began to be scientifically studied in the middle of the nineteenth century (Lynn 2015). In 2015, Dr. Richard Lynn published the first fully comprehensive review on the historical and current evidence of the race differences in intelligence worldwide. In this document, we are interested in examining one of Dr. Lynn’s dataset on the disparity between the intelligence quotient amongst different races, with the direct purpose of discussing its relation to an important quote by Dr. Srinivasan, “I think we should be suspicious when we find ourselves attracted to data—very, very thin and weak data—that seem to justify beliefs that have held great currency in lots of societies throughout history, in a way that is conducive to the oppression of large segments of the population” (Cowen 2021). To conduct this investigation, we use workflow and code similar to Example 2.3. of Chapter 2 of *Telling Stories with Data* (Alexander 2023).

1. Plan

Figure 1 on p. 2 shows a quick sketch of a potential dataset and resulting graph. The dataset should list the important ethnic races, alongside their measured intelligence quotient (Figure 1(a)), and the final graph should display this data as a bar chart (Figure 1(b)).

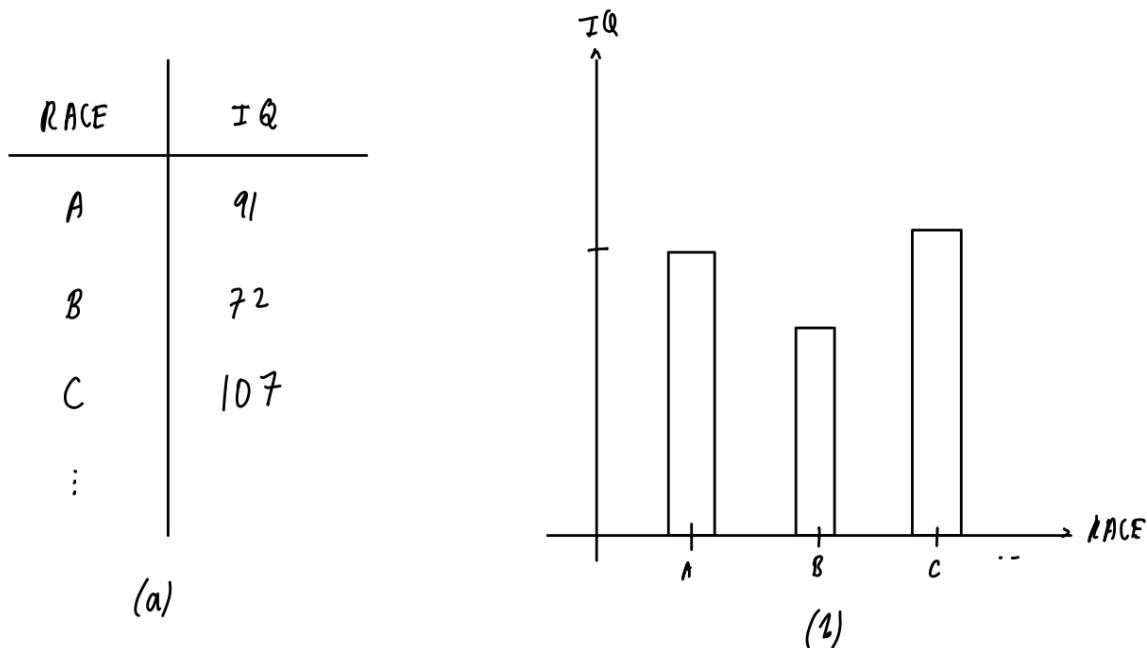


Figure 1: Potential sketch of a dataset and bar graph for a race-based IQ study

2. Simulate

We first set up the workspace by downloading necessary packages in R, as follows.

```
#### Workspace setup ####
#install.packages('tidyverse')
#install.packages('janitor')
#install.packages('formatR')
#install.packages('opendatatoronto')
library(tidyverse)
library(janitor)
library(formatR)
library(opendatatoronto)
knitr::opts_chunk$set(tidy.opts=list(width.cutoff=80), tidy=TRUE) #
↳ prevent source code from running off PDF page
```

To simulate the data, we create a dataset of two variables: **race** and **IQ**. Reasonable values for **race** would be the 11 races to which Dr. Lynn makes reference, these being “Arctic Peoples, Northeast Asians, Europeans, Native Americans, South Asians, North Africans, Bushmen, Sub-Saharan Africans, Australians, Southeast Asians, Pacific Islanders”; reasonable values for

IQ would be randomly sampled from a normal distribution with mean, say, 100, and a standard deviation of 15.

```
#### Simulate ####
set.seed(853) # for reproducibility
simulated_iq_data <- tibble(race = c("Arctic Peoples", "Northeast
  ↳ Asians", "Europeans",
    "Native Americans", "South Asians", "North Africans", "Bushmen",
    ↳ "Sub-Saharan Africans",
    "Australians", "Southeast Asians", "Pacific Islanders"), IQ =
  ↳ round(rnorm(n = 11,
    mean = 100, sd = 15)))

print(simulated_iq_data)
```

```
# A tibble: 11 x 2
  race                IQ
  <chr>              <dbl>
1 Arctic Peoples      95
2 Northeast Asians    99
3 Europeans           73
4 Native Americans    83
5 South Asians        85
6 North Africans     127
7 Bushmen            79
8 Sub-Saharan Africans 93
9 Australians         92
10 Southeast Asians   88
11 Pacific Islanders  125
```

3. Acquire

We use Table 16.2 in Dr. Lynn's *Race Differences in Intelligence: An Evolutionary Analysis* (2015). As the table is not directly downloadable as .csv file, we have copy pasted the data ourselves into an Excel file and thereafter converted it to a .csv file.

```
#### Acquire data ####
raw_iq_race_data <- read_csv(file = "race_iq_data.csv", show_col_types =
  ↳ FALSE)
```

```
#### Save data ####
write_csv(x = raw_iq_race_data, file = "raw_iq_race_data.csv")
```

We may quickly inspect the first six rows of the data set by using `head()` .

```
head(raw_iq_race_data)
```

```
# A tibble: 6 x 5
  RACE          `WINTER TEMP` `WURM TEMP` `BRAIN SIZE`  IQ
  <chr>          <dbl>      <dbl>      <dbl> <dbl>
1 Arctic Peoples      -15        -20        1443    91
2 Northeast Asians    -7         -12        1416   105
3 Europeans           0          -5        1369   100
4 Native Americans     7           5        1366    86
5 South Asians        12           7        1293    84
6 North Africans      12           7        1293    83
```

To make this dataset similar to the one in which we are interested (Figure 1), we will need to change its columns using `clean_names()` and reduce the columns to only the relevant ones using `filter()` .

```
iq_race_data_clean <- clean_names(raw_iq_race_data) |>
  select(race, iq)
head(iq_race_data_clean)
```

```
# A tibble: 6 x 2
  race          iq
  <chr>      <dbl>
1 Arctic Peoples    91
2 Northeast Asians  105
3 Europeans        100
4 Native Americans   86
5 South Asians      84
6 North Africans    83
```

Having appropriately cleaned the dataset, we now save it.

```
write_csv(x = iq_race_data_clean, file = "cleaned_iq_race_data.csv")
```

4. Explore

We explore and visualise the dataset just created by making a bar graph. In Figure 2, we use `ggplot2` of `tidyverse` to build a graph of the IQ of various races.

```
#### Explore data ####
cleaned_iq_race_data <- read_csv(file = "cleaned_iq_race_data.csv",
                                show_col_types = FALSE)
ggplot(cleaned_iq_race_data, aes(x = race, y = iq)) +
  geom_bar(stat = "identity", fill = "skyblue", color = "black") +
  labs(title = "IQ Scores by Race",
       x = "Race",
       y = "IQ Score",
       caption = "Figure 2: Bar graph of IQ scores for various races.")
  ↪ +
  theme_minimal() +
  theme(plot.caption = element_text(hjust = 0.5)) + # centre title
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Rotate
  ↪ x-axis labels for legibility
```

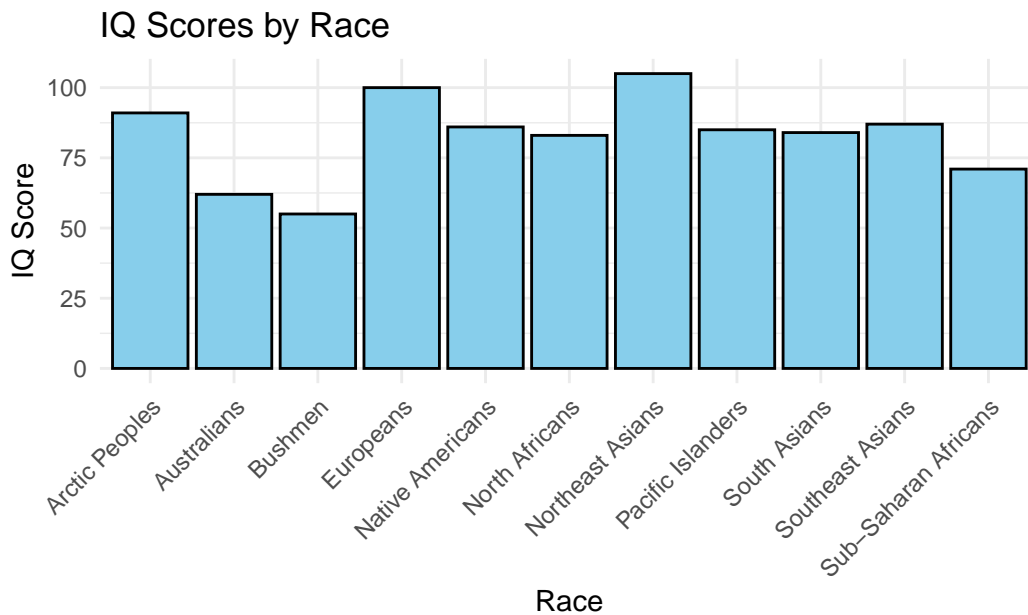


Figure 2: Bar graph of IQ scores for various races.

5. Share

Since the middle years of the nineteenth century, scientific studies concerning race differences in intelligence, and the extent to which they are influenced by genetic and environmental factors, began to emerge (Lynn 2015). The definition of intelligence proposed by the American Psychological Association in 1995, and accepted by Dr. Lynn, is the ability “to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, to overcome obstacles by taking thought” (Neisser 1995). We were interested in visualising the data of the IQ of these eleven categories of races adopted by Dr. Lynn, being Arctic Peoples, Northeast Asians, Europeans, Native Americans, South Asians, North Africans, Bushmen, Sub-Saharan Africans, Australians, Southeast Asians and Pacific Islanders.

We used the data provided in Table 16.2 of Dr. Lynn’s *Race Differences in Intelligence* (2015). We cleaned, tidied and analysed the dataset using the statistical language R (R Core Team 2023) as well as `tidyverse` (Wickham et al. 2019), `janitor` (Firke 2023) and `formatR` (Xie 2023) packages. We then created a histogram of the purported race-based differences in intelligence.

Dr. Lynn’s collected data appears to be consistent with the hypothesis that differences in intelligence amongst races exist. From his data, it is clear that Northeast Asians, Europeans and Arctic Peoples possess, on average, the largest IQs of 105, 100 and 91, respectively, and Sub-Saharan Africans, Australians and Bushmen possess, on average, the least IQs of 71, 62 and 55, respectively. The remaining races can be seen to cluster close together around a mean IQ of 86 and standard deviation of roughly 1.4.

It would appear that these findings justify commonly-held stereotypes and beliefs, as to the differential intelligence and — therefore, in the view of some, worth — of various human races. Consistent with the aforementioned quote by Dr. Srinivasan, this data, that supports the hypothesis of inherent and immutable differences in the intelligence between human races, could contribute to the marginalisation and oppression of certain racial and ethnic groups in modern society. To question the validity of the data, and to scrutinise the means by which it has been acquired and analysed, is thus a necessary question for future study. For instance, it is important to critically examine the selection criteria for the eleven specific ethnic groups chosen for the study, and question whether these are regarded as valid ways across the scientific community of categorising people. In addition, it is important to question the validity and reliability of IQ tests across different cultures and ethnic groups, as cultural bias in testing methods, language proficiency and socio-economic factors may significantly affect IQ scores (Lee 2024). In general, there are various societal and environmental factors, such as access to education, economic opportunities, and healthcare, can significantly impact intelligence test results. A simplistic attribution of intelligence differences to genetics and race alone would overlook these factors.

Another future area of study is to consider the ethical implications of this nature of research,

which suggests intellectual differences between races and, thus, can reinforce harmful stereotypes and contribute to discrimination. In going against principles of fairness, equity and justice, it is crucial to make a thorough investigation into its ethical implications.

For reproducibility, we include a link to this project's [GitHub Repository](#), in which is contained the inputs, outputs and scripts with the directives necessary to reproduce this research.

6. References

- Alexander, Peter. 2023. *Telling Stories with Data with Applications in r*. Chapman & Hall. <https://tellingstorieswithdata.com/>.
- Cowen, Tyler. 2021. "Episode 132: Amia Srinivasan on Utopian Feminism." Conversations with Tyler. <https://conversationswithtyler.com/episodes/amia-srinivasan/>.
- Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. <https://github.com/sfirke/janitor>.
- Lee, Christopher T. 2024. "The Impact of Cultural Factors on IQ Testing." *Medium*. Medium. <https://medium.com/@christophr/the-impact-of-cultural-factors-on-iq-testing-be597d33d0f9>.
- Lynn, Richard. 2015. *Race Differences in Intelligence: An Evolutionary Analysis*. Washington Summit Publishers.
- Neisser, Ulric. 1995. *Intelligence: Knowns and Unknowns*. The American Psychological Association.
- R Core Team. 2023. "R: A Language and Environment for Statistical Computing." <https://www.R-project.org/>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Golemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.
- Xie, Yihui. 2023. *formatR: Format r Code Automatically*. <https://github.com/yihui/formatR>.