How Diverse is the MoMA: While Modern, Not Diverse in Its Artists and Directors*

An analysis of the MoMA's public datasets on their collections and exhibits

Sakura Ariga

17 April 2023

The MoMA's public GitHub data on their exhibits and collections was analyzed to determine how diverse the artists and directors at the MoMA are. It was found that the MoMA is comprised largely of white men, but that this relationship has changed with time. This finding gives quantitative proof of the need for more representation within not jus the MoMA, but all American museums as a whole.

1 Introduction

Diversity in museums is a topic that is increasingly being discussed and raised as a problem in museums. While the main goal of museums is to serve as cultural institutions and a place of learning, oftentimes the objects on display or the decision-making staff who work behind the scenes do not meet this mission of inclusion.

This paper will investigate the level of diversity in one such museum: the Museum of Modern Art (also known as MoMA). Founded in 1929, the MoMA, as its name suggests, houses a collection of contemporary art pieces. It was originally created to go against traditional art museums by housing only modern art, and has since grown to become an influential cultural institution. Using data made publicly available on the MoMA's GitHub account, this paper aims to evaluate how diverse the artists of the artworks on display are, as well as how diverse the directors and department heads who chose to display these artworks are, in terms of gender and nationality. Thus, the estimand is: how much of the artists and directors in the MoMA's history are/were not white men.

It was found that the MoMA's artists and directors were predominantly men. A simple linear regression model shows that time affects this proportion of men.

*Code and data are available at: https://github.com/sakura-ariga/MOMA_diversity_analysis.git.

The paper's structure begins by explaining the data used, followed by a linear model, a discussion of the relationship between variables in the data and in the model, and a discussion on key findings.

2 Data

Investigating artists:

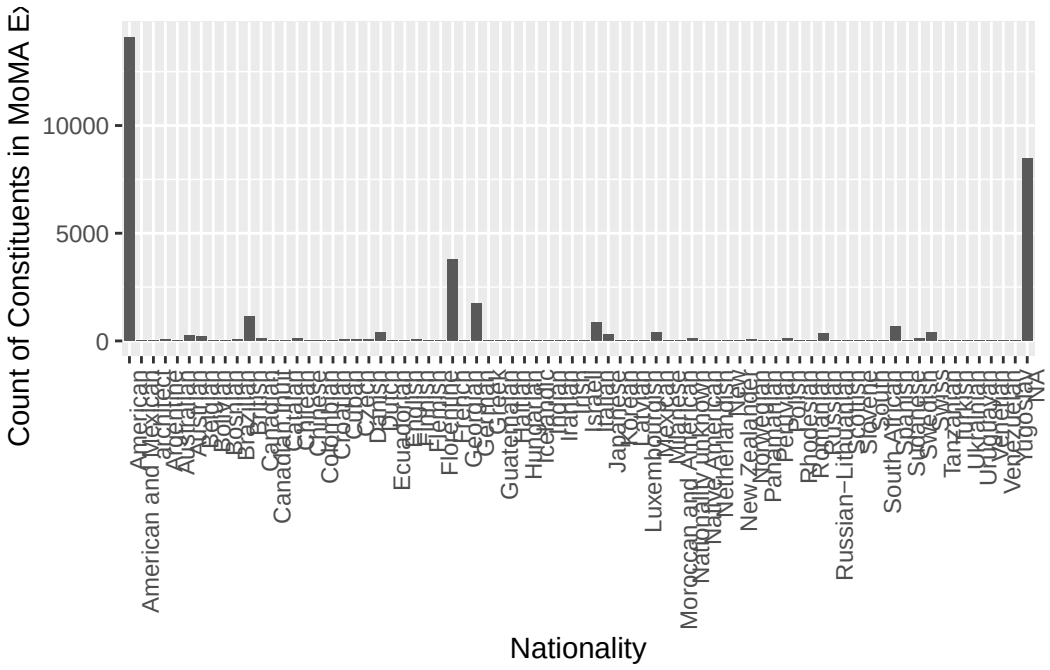

Figure 1: Nationality Distribution of Consituents (Artists, Staff) in MoMA Exhibits

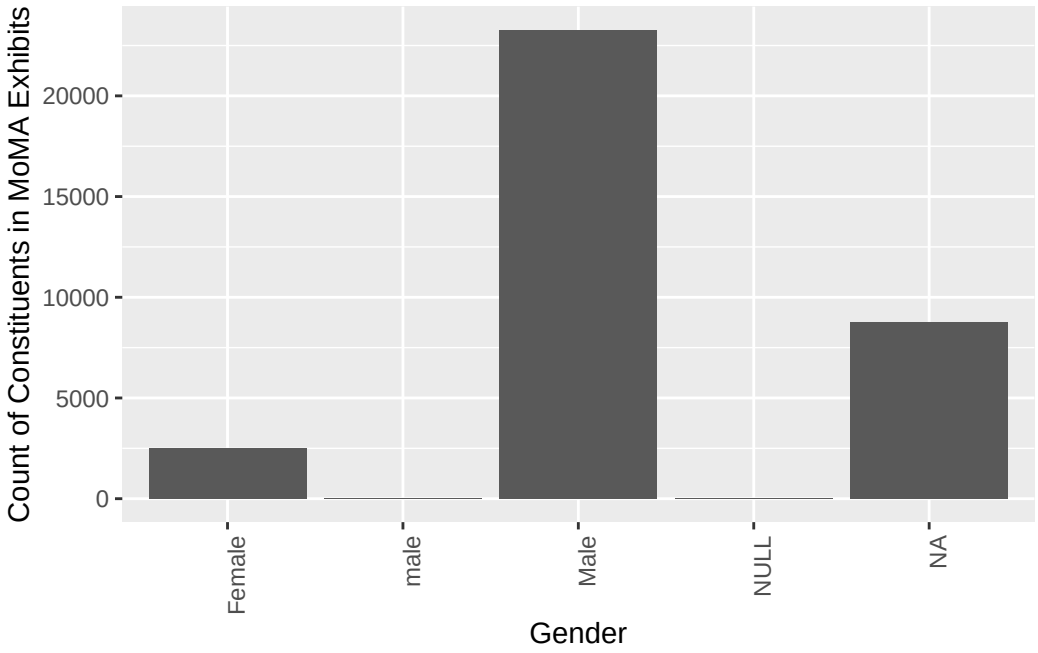

Figure 2: Gender Distribution of Consituents (Artists, Staff) in MoMA Exhibits

Investigating directors and department heads:

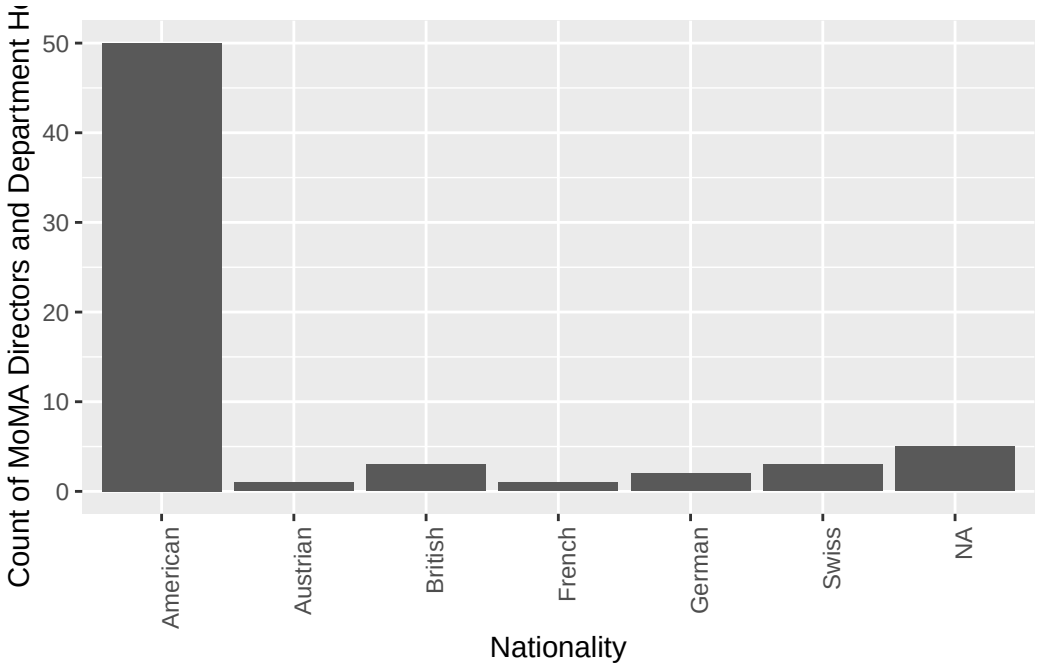

Figure 3: Nationality Distribution of Directors and Department Heads in MoMA

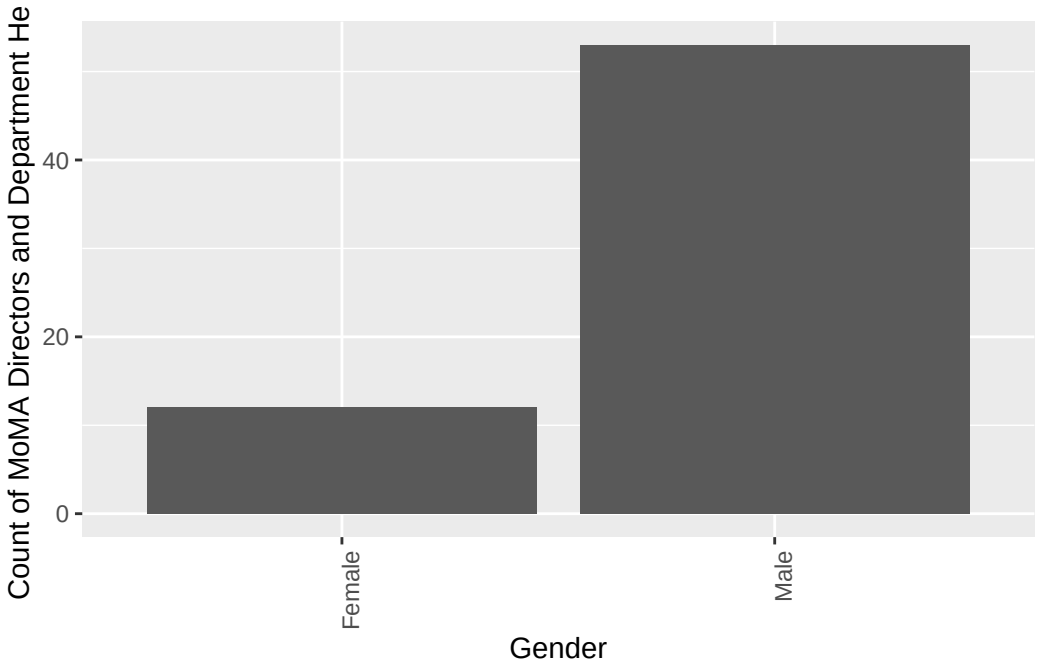

Figure 4: Gender Distribution of Directors and Department Heads in MoMA

3 Model

```
Call:
lm(formula = Gender ~ DepartmentBeginYear, data = numeric_momadirectors_data)

Residuals:
    Min      1Q  Median      3Q     Max
-0.8440  0.1560  0.1608  0.1882  0.2912

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)         4.948882   4.073112   1.215    0.229
DepartmentBeginYear -0.001610   0.002092  -0.769    0.445

Residual standard error: 0.3923 on 63 degrees of freedom
Multiple R-squared:  0.009308,  Adjusted R-squared:  -0.006417
F-statistic: 0.5919 on 1 and 63 DF,  p-value: 0.4446
```

The goal of our modelling strategy is twofold. Firstly,...

Here we briefly describe the Bayesian analysis model used to investigate... Background details and diagnostics are included in Appendix B.

3.1 Model set-up

Define y_i as the number of seconds that the plane remained aloft. Then β_i is the wing width and γ_i is the wing length, both measured in millimeters.

$$y_i|\mu_i, \sigma \sim \text{Normal}(\mu_i, \sigma) \tag{1}$$
$$\mu_i = \alpha + \beta_i + \gamma_i \tag{2}$$
$$\alpha \sim \text{Normal}(0, 2.5) \tag{3}$$
$$\beta \sim \text{Normal}(0, 2.5) \tag{4}$$
$$\gamma \sim \text{Normal}(0, 2.5) \tag{5}$$
$$\sigma \sim \text{Exponential}(1) \tag{6}$$

We run the model in R (**citeR?**) using the `rstanarm` package of (**rstanarm?**). We use the default priors from `rstanarm`.

Table 1: Explanatory models of flight time based on wing width and wing length

	First model
(Intercept)	4.87
	(4.04)
DepartmentBeginYear	0.00
	(0.00)
Num.Obs.	65
R2	0.012
R2 Adj.	−0.040
Log.Lik.	−30.994
ELPD	−33.7
ELPD s.e.	6.7
LOOIC	67.4
LOOIC s.e.	13.3
WAIC	67.3
RMSE	0.39

3.1.1 Model justification

We expect a positive relationship between the size of the wings and time spent aloft. In particular…

We can use maths by including latex between dollar signs, for instance θ.

4 Results

Our results are summarized in Table 1.

5 Discussion

5.1 Gender distribition

5.2 Nationality distribition

5.3 Analyzing the model: is time a factor?

5.4 Comparing artists vs directors

5.5 Weaknesses and next steps

Appendix

A Additional data details

B Model details

B.1 Posterior predictive check

In **?@fig-ppcheckandposteriorvsprior-1** we implement a posterior predictive check. This shows…

In **?@fig-ppcheckandposteriorvsprior-2** we compare the posterior with the prior. This shows…

Examining how the model fits, and is affected by, the data

Figure 5: **?(caption)**

B.2 Diagnostics

?@fig-stanareyouokay-1 is a trace plot. It shows… This suggests…

?@fig-stanareyouokay-2 is a Rhat plot. It shows… This suggests…

Checking the convergence of the MCMC algorithm

Figure 6: **?(caption)**

C References