# Jean Pockets and Gender*

**subtitle**

Hari Lee Robledo

April 23, 2024

Fashion has long reflected and reinforced gender inequalities, with women's clothing often lacking functional pockets found in men's garments. A 2018 study by Jan Diehm & Amber Thomas on The Pudding revealed this discrepancy, highlighting inferior pocket functionality in women's jeans. This paper aims to reproduce these findings and explore if gender disparities extend to pricing. Using the same data, we analyze the correlation between prices of men's and women's jeans. Results suggest that women's jeans not only have smaller pockets but also tend to be more expensive. This study sheds light on gender inequalities in everyday products and their broader implications. [UPDATE according to findings]

## Table of contents



*Code and data are available at: https://github.com/hari-lr/pockets-and-gender

# 1 Introduction

Throughout history, women have faced numerous restrictions and limitations, often ingrained within societal norms and cultural practices. Among these limitations, the realm of fashion has served as both a reflection and an enforcer of gender-based inequalities. From ancient times to the modern era, clothing has been a means of expressing social status, cultural identity, and gender roles. According to design expert, Hannah Carlson, pocket inequality stems from the way clothes are designed and made, she explains that "From a very early age, I think we sort of agree as a culture that womenswear, girlswear is meant to be pretty," she said. "And menswear, boyswear is meant to be utilitarian" (Alwahaidi 2018). Men's garments historically incorporated functional pockets, allowing them to carry essentials such as money, keys, and tools, women's clothing has often lacked this practical feature. Instead, women's garments have frequently been designed with form-fitting silhouettes, delicate fabrics, and ornamental details, prioritizing aesthetics over utility.However, the 1880s, women began to be vocal about pocket inequality, as the "demand for the vote and the demand for pockets were made together" (Alwahaidi 2018).

In 2018, Jan Diehm & Amber Thomas published Womens´s Pockets are Inferior on *The Pudding*, a data-centric digital publication. In this paper, they "measured pockets in both men´s and women's pants in 20 of the US' most popular blue jeans brands" (Jan Diehm 2018). They "programmatically determined whether various everyday items could fit in an otherwise empty pocket in jeans that aren't being worn" (Jan Diehm 2018) and concluded that women´s pockets were less functional than men´s.

Using the data from the previously described article, this paper will graphically reproduce the results and examine whether there is a correlation between the prices of men's jeans and women's jeans. The study seeks to determine if women's jeans not only have smaller and less functional pockets but also tend to be more expensive. This will help better understand if gender inequality is only reflected in the jean's pockets or in pricing as well. The findings revealed that pocket area alone may not be a strong predictor of jeans prices for either gender, and there may be other factors influencing the pricing of jeans. A linear regression analysis revealed divergent findings regarding the relationship between pocket area and price in men's and women's jeans. For men's jeans, no significant correlation was found between pocket area and price, indicating that changes in pocket size do not significantly affect the price of men's jeans. In contrast, although a negative correlation was observed in women's jeans, it was not statistically significant, suggesting that while larger pocket areas might be associated with slightly lower prices, other factors likely play a more influential role in determining the price of women's jeans.

This paper is structured into the following sections: Data, which explains the collection and cleaning process; Results, presenting trends and correlations found in the data; Discussion, comparing and evaluating the data; and Conclusion, summarizing the findings.

Table 1: **?(caption)**

```
# A tibble: 5 x 64
  brand           style menWomen name  fabric price maxHeightFront minHeightFront
  <chr>           <chr> <chr>    <chr> <chr>  <dbl>          <dbl>          <dbl>
1 7 for All Man~ stra~ men      The ~ 98% c~ 179              21.5             23
2 7 for All Man~ skin~ men      Paxt~ 92% c~ 209              21               23
3 Abercrombie    slim  men      Lang~ 95% c~  78              22               15
4 Abercrombie    stra~ men      Kenn~ 95% c~  78              22.5             15
5 American Eagle stra~ men      Stra~ 85% c~  50.0            22               18
# i 56 more variables: rivetHeightFront <dbl>, maxWidthFront <dbl>,
#   minWidthFront <dbl>, maxHeightBack <dbl>, minHeightBack <dbl>,
#   maxWidthBack <dbl>, minWidthBack <dbl>, cutout <lgl>, waistSize <dbl>,
#   updatedStyle <chr>, group <chr>, priceGroup <chr>, pocketArea <dbl>,
#   rectanglePhone.area <dbl>, rectanglePhone.cx <dbl>,
#   rectanglePhone.cy <dbl>, rectanglePhone.width <dbl>,
#   rectanglePhone.height <dbl>, rectanglePhone.angle <dbl>, ...
```

Table 2: Sample of Cleaned Pocket Area, Price and Gender Data

| brand | men_women | price | pocket_area |
|---|---|---:|---:|
| 7 for All Mankind | men | 179.00 | 11222.73 |
| 7 for All Mankind | men | 209.00 | 10841.94 |
| Abercrombie | men | 78.00 | 10624.65 |
| Abercrombie | men | 78.00 | 10134.15 |
| American Eagle | men | 49.95 | 10555.40 |

## 2 Data

The data utilized in this paper was retrieved from The Pudding GitHub Site Portal, specifically the data collected by by Jan Diehm and Amber Thomas [CITE]. Data was collected, cleaned, and analyzed using the open-source statistical programming software R (R Core Team 2023). This process involved various packages within R, including tidyverse (Wickham et al. 2019), `ggplot2` (**ggplot2?**), `dplyr` (**dplyr?**), `readxl` (**readxl?**), `tibble` (**tibble?**), `janitor` (**janitor?**), `KableExtra` (**kableExtra?**), `knitr` (**knitr?**), `ggbeeswarm` (**ggbeeswarm?**), and `ggrepel` (**ggrepel?**). A comprehensive description of the data gathering and cleaning process is provided in the following subsections. [MAKE SURE ALL LIBRARIES ARE CITED]

### 2.1 JSON Dataset and cleaning

## 3 Model

Given the similarity in lightness observed in foundation shades across both countries, it was of interest to investigate if any differences in saturation existed. This investigation was motivated by the fact that the local population in each country has unique undertones. Before conducting the analysis, several tests were conducted to verify that the model assumptions (See Appendix A), including linearity, normality, and homoscedasticity of residuals, were satisfied, thereby ensuring that the model was well-suited for the data at hand.

The null hypothesis: there is no significant linear relationship between the lightness of foundation shades and the saturation of undertones.

The alternative hypothesis: there is a significant linear relationship between the lightness of foundation shades and the saturation of undertones.

A linear regression analysis can be conducted to test this hypothesis and to determine whether a statistically significant relationship exists between the variables. A low p-value for the regression coefficient for saturation would provide evidence against the null hypothesis, indicating that changes in saturation do result in a meaningful variation in lightness across foundation shades. On the other hand, if the p-value is high, this would suggest that there is not enough evidence to reject the null hypothesis, and there may be no significant linear relationship between the variables. Ultimately, the results of the analysis will inform whether the null hypothesis can be rejected or not. The mathematical equation for the linear regression model can be written as:

$$L = \beta_0 + \beta_1 S + \epsilon$$

Where:

- $L$ represents the dependent variable, which is the overall lightness of the foundation shade
- $S$ represents the independent variable, which is the saturation of the foundation shade
- $\beta_0$ represents the intercept of the regression line, which is the predicted value of $L$ when $S$ is equal to zero
- $\beta_1$ represents the slope of the regression line, which is the change in $L$ for a one-unit increase in $S$
- $\epsilon$ represents the random error term, which accounts for variability in $L$ that is not explained by the relationship with $S$

The aim of linear regression is to estimate the values of $\beta_0$ and $\beta_1$ that minimize the sum of squared errors between the predicted and actual values of lightness ($L$) for the given values of saturation ($S$) in the dataset. By minimizing this sum of squared errors, the linear regression model can produce the best-fit line to describe the linear relationship between the two variables, allowing for the estimation of $L$ values based on known $S$ values.

In this project, the lightness variable ($L$) is measured on a scale of 0 to 100, while the saturation variable ($S$) is measured on a scale of 0 to 1. Therefore, the slope coefficient $\beta_1$ in the linear regression model would represent the change in lightness ($L$) for a one-unit increase in saturation ($S$), where a one-unit increase in $S$ corresponds to an increase of 1/100 in the saturation of the foundation shade. The intercept $\beta_0$ in the model would represent the expected value of $L$ when $S$ is equal to zero, which in this case would correspond to the lightness value for a completely desaturated foundation shade.

# 4 Results

```
Call:
lm(formula = price ~ pocket_area, data = men_jeans_data)

Residuals:
   Min     1Q Median     3Q    Max 
-47.33 -28.80 -10.43  12.97 127.88 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)  
(Intercept) 105.900217  53.525444   1.979   0.0552 .
pocket_area  -0.002286   0.004913  -0.465   0.6445  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 40.9 on 38 degrees of freedom
Multiple R-squared:  0.005662,	Adjusted R-squared:  -0.0205 
F-statistic: 0.2164 on 1 and 38 DF,  p-value: 0.6445
```

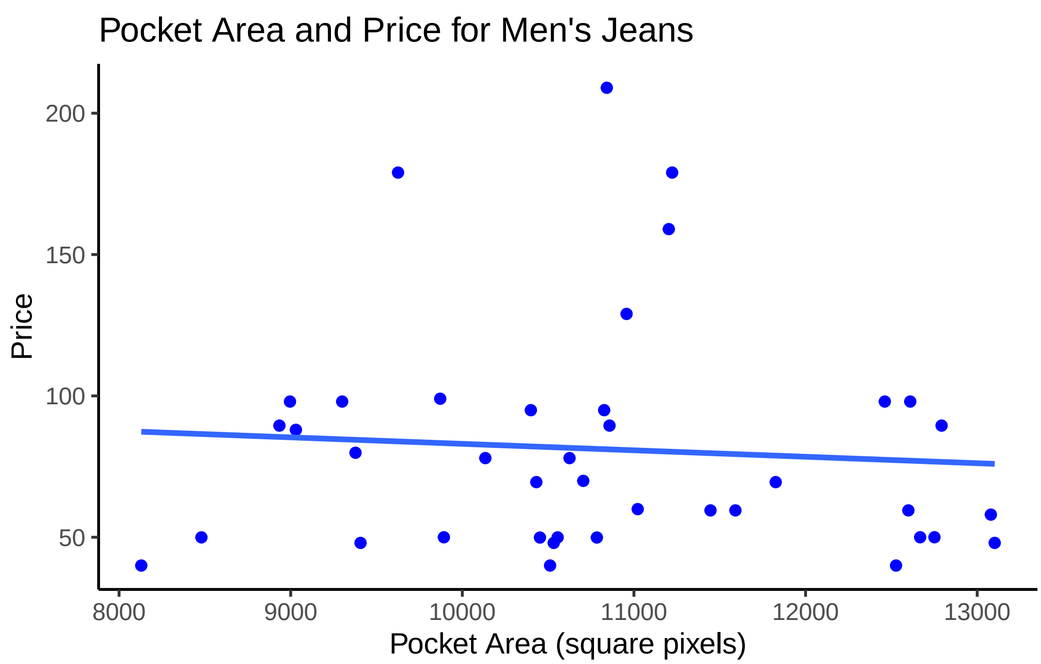


Figure 1: Jean Price vs Pocket Size (Men)

```
Call:
```

```
lm(formula = price ~ pocket_area, data = women_jeans_data)

Residuals:
    Min      1Q  Median      3Q     Max
-59.980 -32.199  -3.117  16.964 162.600

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 136.715960  42.032486   3.253   0.0024 **
pocket_area  -0.009345   0.006848  -1.364   0.1804
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 48.26 on 38 degrees of freedom
Multiple R-squared:  0.04671,   Adjusted R-squared:  0.02162
F-statistic: 1.862 on 1 and 38 DF,  p-value: 0.1804
```

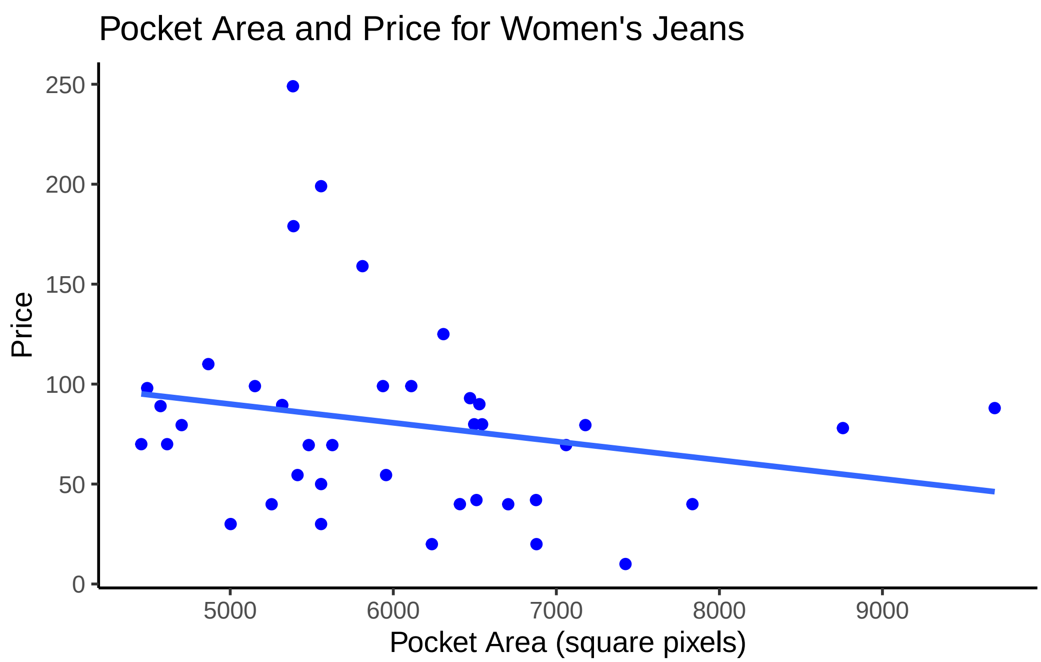


Figure 2: Jean Price vs Pocket Size (Women)

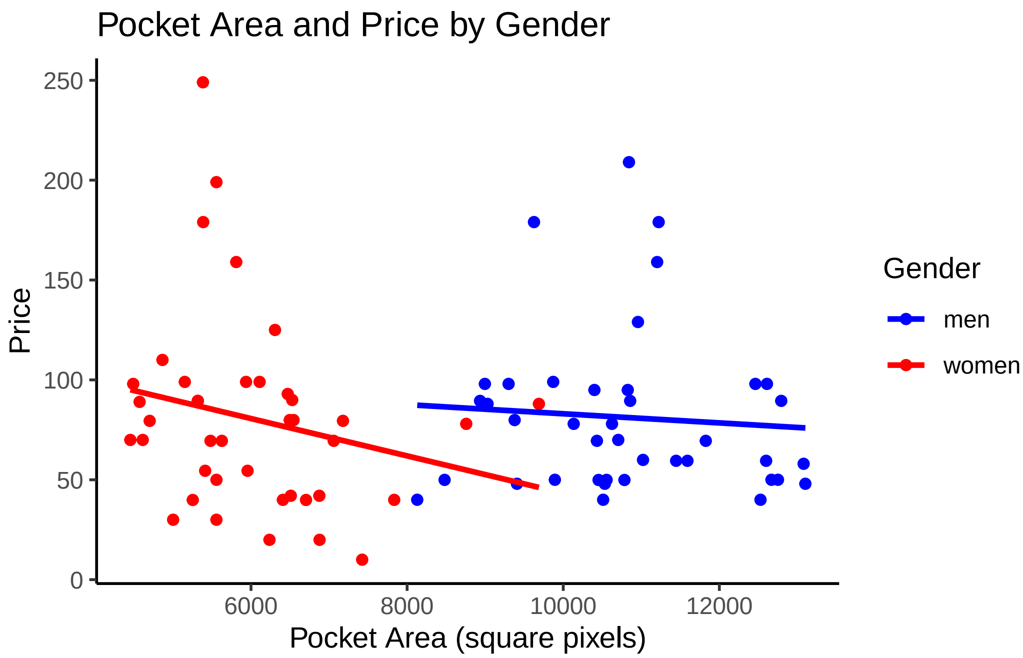


Figure 3: Jean Price vs. Pocket Area Size per Gender

# 5 Discussion

## 5.1 First discussion point

## 5.2 Second discussion point

## 5.3 Third discussion point

##Weaknesses and next steps

## Appendix

## References


Alwahaidi, Keena. 2018. *Women's Apparel Doesn't Have Enough Pockets. This Expert Says That Has to Change.*

Jan Diehm, Amber Thomas. 2018. *Womens´s Pockets Are Inferior.*

R Core Team. 2023. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.