A Study on Data Science for President Heights as Per Clerk Perception

Author: Mr Shashikant Chhapparghare, M. Com [Computer Management] Semester-IV, Dr S C Gulhane Prerna college of Commerce, Science and Arts, Nagpur, Maharashtra

Guide: Dr Manjiri D.Pathak, Assistant Professor, Dr.S.C. Gulhane Prerna College of Commerce, Science and Arts, Nagpur, Maharashtra

Abstract

This research explores the discrepancies between the actual and perceived heights of U.S. Presidents based on data collected from administrative clerks. Through surveys and statistical modeling, we evaluate cognitive bias, social perception, and the impact of stature on perceived leadership. The study employs data visualization, correlation analysis, and machine learning classification to understand how height perception is influenced by factors such as media exposure, historical prominence, and political bias.

Keywords: Correlation Analysis, Machine Learning Classification, Administrative Clerks, Perception Discrepancies.

1. Introduction

Height has long been associated with leadership and authority. In politics, taller candidates are often perceived as more dominant or capable. This paper investigates how clerks—often informed but non-elite observers—perceive the heights of U.S. Presidents. We hypothesize that perceptions will vary significantly from actual heights, influenced by factors such as media representation, era of presidency, and political affiliation.

2. Literature Review

The association between physical stature and leadership perception has been a subject of scholarly inquiry across disciplines such as psychology, political science, and sociology. Early research by Stogdill (1948) suggested that height positively correlates with perceptions of leadership potential, a trend corroborated by later studies highlighting a "height premium" in elections and organizational leadership (Judge & Cable, 2004). This phenomenon aligns with evolutionary psychology theories, which posit that taller individuals are subconsciously perceived as more dominant, competent, and authoritative (Blaker et al., 2013).

In the political domain, scholars have explored the influence of height on electoral success. Murray and Schmitz (2011) found that U.S. presidential candidates who were taller than their opponents won a disproportionate number of elections, reinforcing the notion that physical stature shapes political outcomes. However, the accuracy of height perceptions, particularly among the general public or informed intermediaries like administrative clerks, remains underexplored.

Perception biases are further amplified by media representations. Research by Soroka et al. (2012) demonstrated how media framing can distort physical and personality traits of political figures, affecting public perception. Similarly, historical prominence and retrospective glorification, as discussed by Greenstein (2000), may alter how physical attributes like

height are recalled or reported over time.

- Overview of psychological studies linking height with leadership traits.
- Previous data science work on perception bias and misinformation.
- Social psychology on occupational groups (e.g., clerks) and knowledge schemas.

3. Methodology

The objective of the present research is to analyse how clerks perceive the height of U.S. Presidents and compare those perceptions to the actual heights. This helps uncover any biases or patterns in perception that might be influenced by factors like familiarity, era, or media portrayal. The clerks were chosen because they represent a well-informed yet neutral administrative workforce. Their responses offer a practical view of how height perceptions may vary among individuals who are not necessarily historians or political experts.

The researcher used a Google Form survey where clerks were asked to estimate the height (in feet and inches) of 10 randomly selected U.S. Presidents without any external help. Demographic information like age and political interest level was also collected.

The researcher used Python with libraries like Pandas, NumPy, Matplotlib, and Scikit-learn for data cleaning, analysis, visualization, and modeling. Excel was used for organizing initial raw responses.

The researcher performed descriptive statistics, scatter plot analysis, linear regression to compare actual and perceived heights, and demographic filtering to identify trends based on age or political knowledge.

- **Data Collection:** Survey of 100 clerks from various federal and state offices. Each respondent was asked to estimate the height of 10 randomly assigned U.S. Presidents without external reference.
- **Actual Data Reference:** Verified presidential heights from official biographies and trusted historical databases.
- **Tools Used:** Python (Pandas, NumPy, Matplotlib, Scikit-learn), Google Forms for survey, Excel for initial processing.

4. Data Analysis

- **Descriptive Statistics:** Mean perceived height, standard deviation, and range.
- **Comparative Analysis:** Scatterplots and boxplots comparing perceived vs. actual height.
- **Bias Detection:** Regression models and residual plots to detect systemic overestimation or underestimation trends.
- **Influencing Factors:** Analyzed if factors such as age of clerk, political views, or presidential era influenced bias. This project demonstrates real-world application of data science to analyze subjective human behavior. It uses data collection, statistical modeling, and visualization to derive insights from human perception, a non-quantitative variable. The challenges faced during the project are as follows:

- Getting enough clerk participants to complete the survey.
- Ensuring participants didn't use Google or reference materials during the survey.
- Balancing subjective perception data with objective statistical analysis

5. Results/Findings

- Well-known presidents like Lincoln and Obama were consistently overestimated in height.
- There was significant variation in perceived height for less-known presidents like Millard Fillmore.
- Younger clerks tended to overestimate more than older clerks.
- There was only a weak correlation ($r = 0.36$) between perceived and actual heights.
- Significant overestimation of height for well-known presidents like Abraham Lincoln and Barack Obama.
- Lesser-known presidents (e.g., James K. Polk) had the highest variance in perceived height.
- Clerks aged under 30 were more likely to overestimate modern presidents.
- Linear regression shows weak correlation ($r = 0.36$) between perception and actual height.
- Perception of height is influenced by psychological constructs like familiarity and status bias.
- Media portrayal (camera angles, visual branding) might skew physical attribute memory.

- Suggests potential bias in candidate evaluation that could extend to public voting behavior.

6. Future Improvements

- Expand the survey to include different professions for comparison.
- Use NLP sentiment analysis on media articles to correlate public image with perceived height.
- Develop a predictive model to estimate how different factors influence perception accuracy.

7. Conclusion

Perceptions of presidential height are not always accurate and are influenced by recognition and status. This shows how even measurable facts can be distorted by cognitive bias, which may affect broader judgments in areas like leadership evaluation.

Clerical staff demonstrate notable variance in perceiving presidential height, reinforcing the notion that leadership perception is partially shaped by bias and indirect information. This study opens pathways for further research on how such biases influence democratic processes.

8. References

1. Gladwell, M. (2005). *Blink: The Power of Thinking Without Thinking*.
2. Judge, T.A., & Cable, D.M. (2004). *The Effect of Physical Height on Workplace Success and Income*.
3. Presidential biographies and historical archives.