



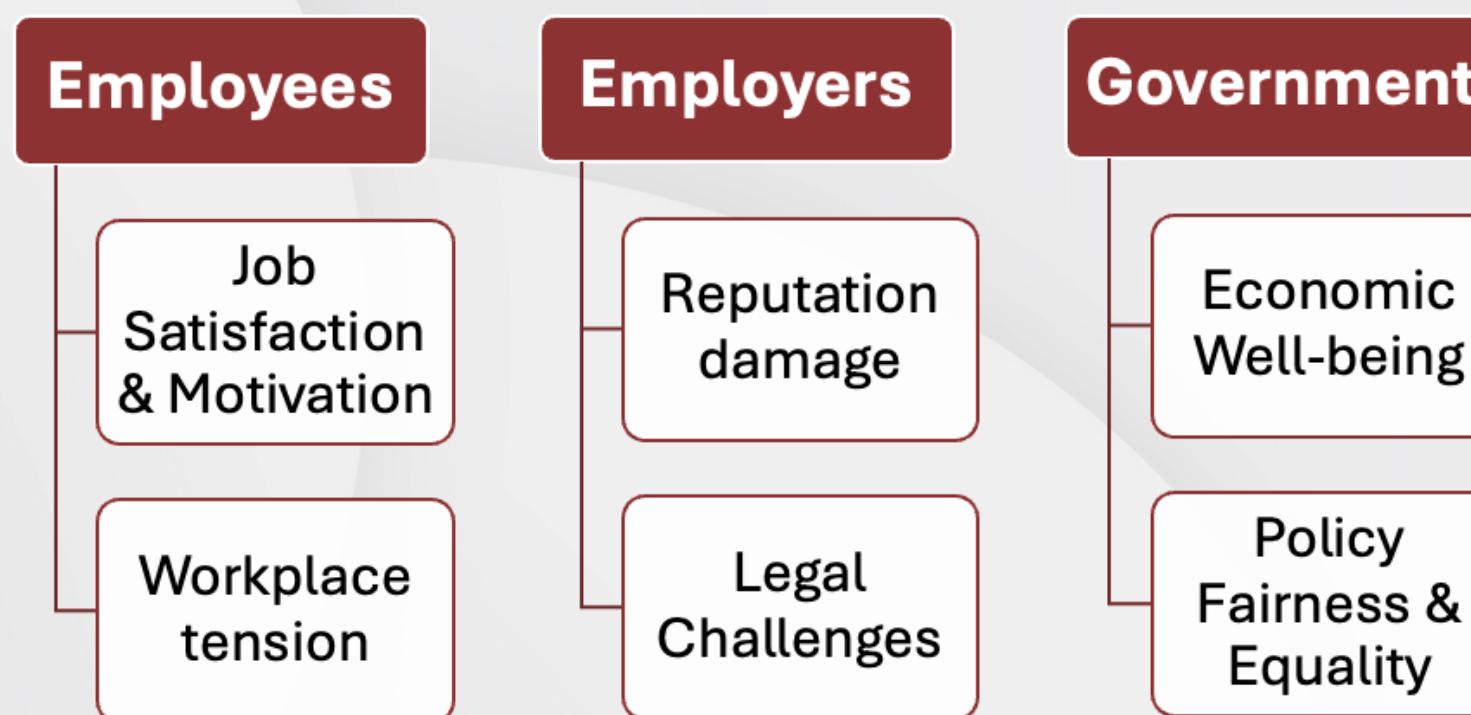
GENDER-WAGE GAP ANALYSIS

BY TEAM 33

SNOWY TO, BINGO UPATISING, JOEY CINDASS, ROXANNE LI

GENERAL BACKGROUND

- WHY should we care?

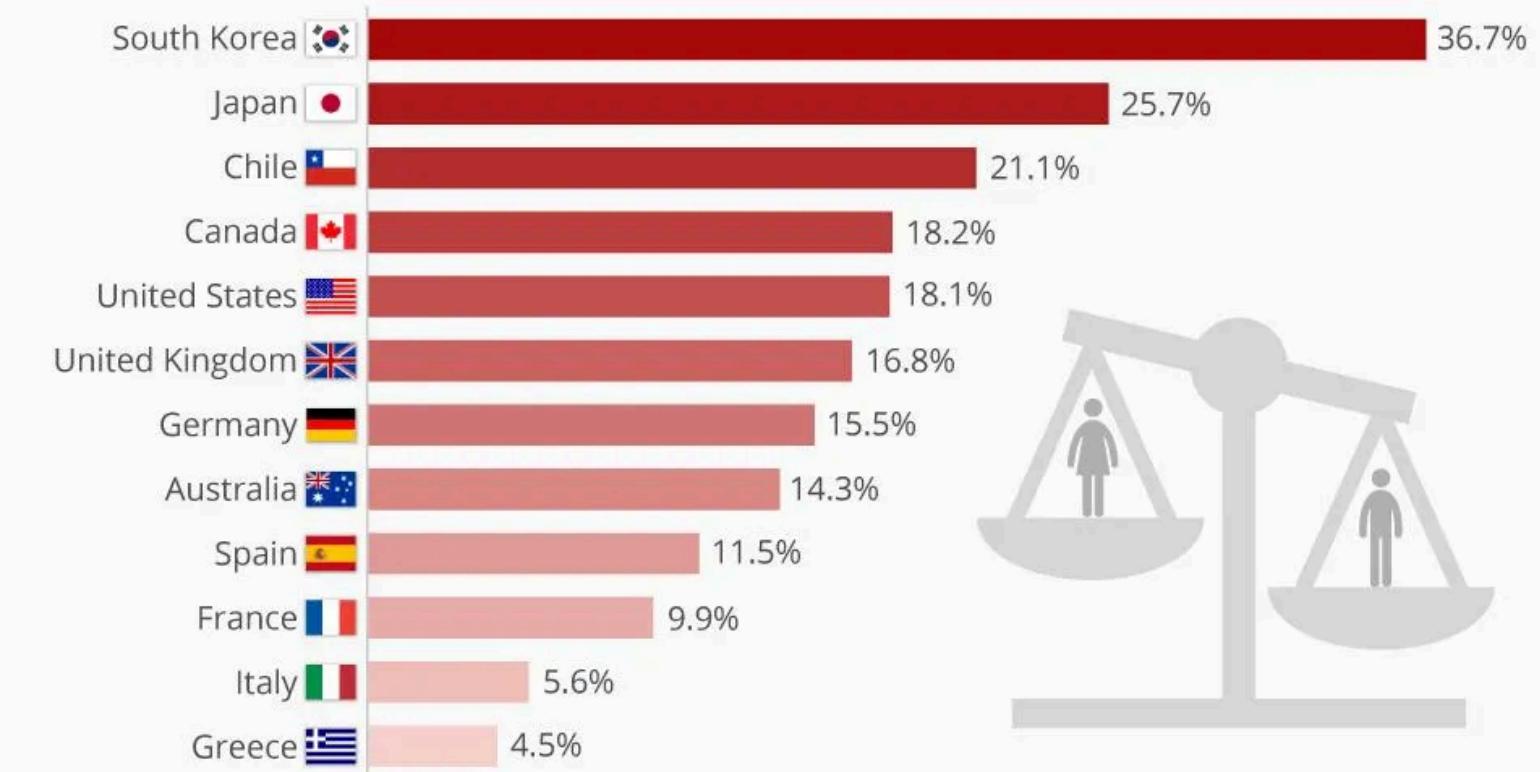


- WHY wage disparity happens?

- Main contributors
- Actionable insights

Where The Gender Pay Gap Is Widest

% difference in full-time earnings between men/women in selected OECD nations*



* As a % of the earnings of men, 2016 or latest available year.
Source: OECD

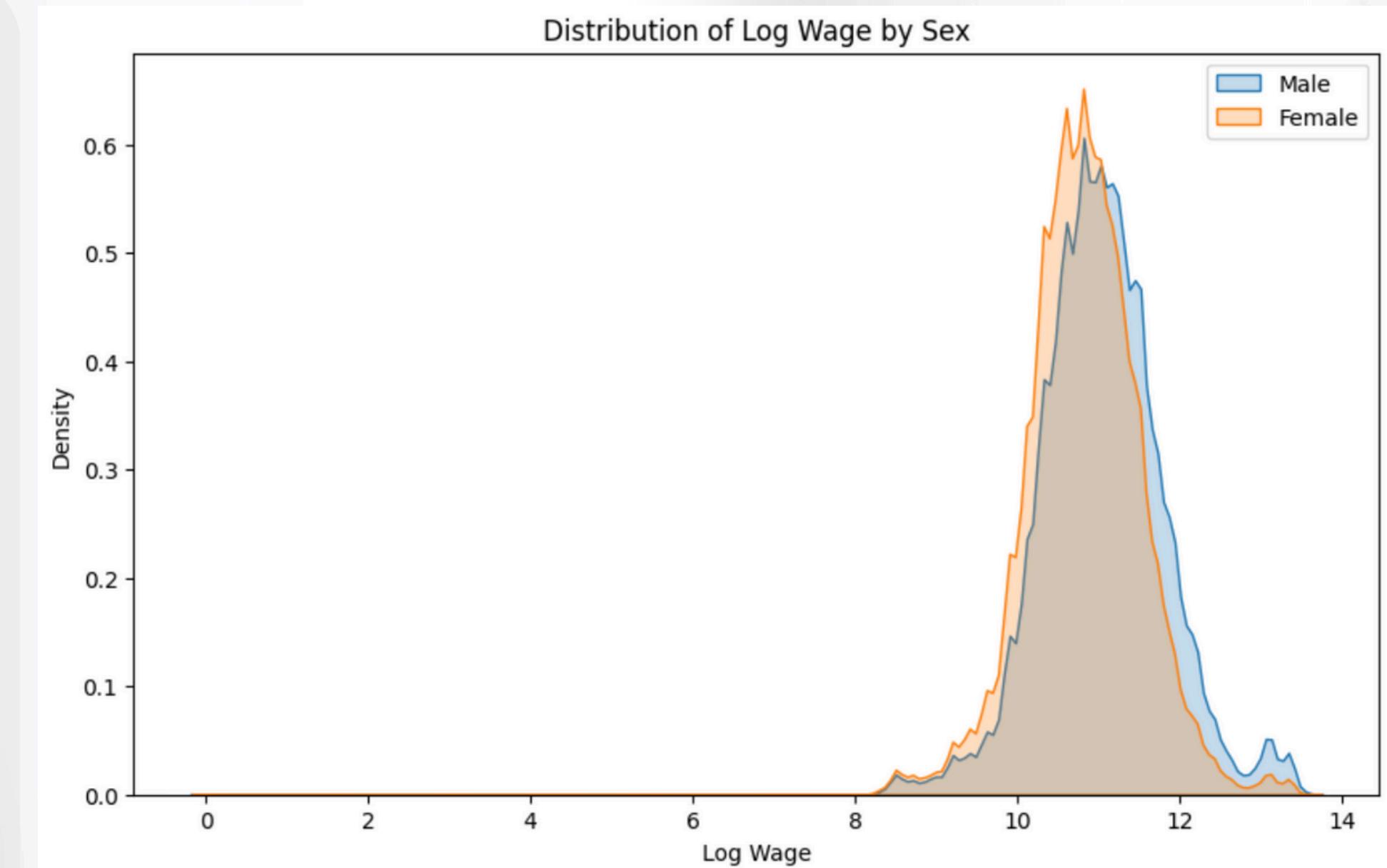


Forbes statista



OUR DATA

- **Data Source**
 - 2021 U.S. Census
- **Data Cleaning**
 - Dropped Columns:
 - NaN, duplicate values, irrelevance
 - [Weight, Degree Fields, Citizenship, Marriage, Ancestor, Employment Status, Income, Investment, Depart, Arrives]
 - Kept Columns
 - Mapped [OCCSOC] column into actual Occupation Sectors
 - Added dummy variables (e.g. Male/Female, Occupation)



METHODOLOGY

- Transformation of the target variable Wage
- Linear Regression
 - $y = \log(\text{Wage})$
 - $x = \text{SEX}$
 - Statistically significant; R_square (~2.2%)
- Other Regression Model
 - $\log(\text{Wage}) \sim \text{SEX} * \text{occupation}$ (Statistically significant; R_square (~1.8%))
 - $\log(\text{Wage}) \sim \text{SEX} * \text{AGE}$ (Statistically significant; R_square (~2.0%))
- Converted objects into INT using label and one-hot encoding
- Using Lasso for variable selection and regularization by turning some coefficients to zero, improving model simplicity and preventing overfitting
- Use chosen variables to fit the regression
 - Statistically significant; R_square (~4%)



INSIGHTS

After controlling for factors like occupation, education, work hours, and state, being male is associated with ~\$19,100 higher income compared to being female.

Feature	Coefficient	Insight
OCC_management	+78,452	Management jobs earn the most on average.
OCC_legal	+56,883	Legal occupations are high-paying.
STATEFIP_District of Columbia	+31,442	Strong regional income effect — high wages in D.C.
OCC_healthcare	+28,961	Healthcare jobs also earn significantly more.
STATEFIP_Connecticut	+26,296	High average wages in this state too.



INSIGHTS

Feature	Coefficient	Insight
AGE_GROUP_young_adult	-21,815	Younger adults (18–29) earn much less than adults (30–62).
OCC_farming_fishing_forestry	-19,359	Agricultural-related work is significantly underpaid.
OCC_edu	-18,072	Education-related occupations tend to pay less.
AGE_GROUP_teen	-4,218	Teenagers earn much less (as expected).
OCC_personal_care_service_occ	-9,558	Low wages in personal care roles.



ACTIONABLE SUGGESTIONS

Insight	Suggestions
Occupation	<ul style="list-style-type: none">• Employees<ul style="list-style-type: none">◦ Personal skill development◦ Consideration of industry switch• Employers<ul style="list-style-type: none">◦ Talent acquisition strategy; training◦ Equitable hiring and promotion practices
Location	<ul style="list-style-type: none">• Employees<ul style="list-style-type: none">◦ Relocation/Remote opportunities• Employers<ul style="list-style-type: none">◦ Location-Based Wage Adjustments



LIMITATIONS

- **Data Generation Process**
 - Data not collected for business analytics purpose
 - Could be other variables of significant impact, but not captured by census
 - Data input by individuals, could be biased/human errors
 - Ex. Ancestors, Language
- **Assumptions in Model Building**
 - Fail to capture nonlinear relationships
 - Potentially omit variables that correlate with others
 - Our data cleaning, Ex. Classification of Age Groups, might be biased.



CONCLUSION

- **Significant Difference in Male/Female Wage, but...**
 - Occupation Sectors
 - Location
 - Management Level
- **In the Future...**
 - Promote wage transparency
 - Using data to detect and eliminate biases
 - Improve on data generation process



Q & A



THANK YOU!