ECON0150 Final Project – Charley Wan & Asliddin Nurboev

The gender pay gap has persisted since the 1860s and with new jobs emerging from industries, trying to equalize the pay gap has become harder. In this project, we were curious if the overall gender pay gap changes once we narrow it down to specific job titles.

Question: *Does the gender wage gap change when we control for Job Title?*

Null: *There is no significant difference in average 'BasePay' between genders after controlling for 'JobTitle'.*

Although there may still be a general disparity between what a man earns and what a woman earns, specific job types may have made more progress in achieving a smaller or zero pay gap.

Summary Statistics & Data Visualizations

Dataset: Glassdoor Salary Statistics on Kaggle (data coverage until 02/28/2019)

Glassdoor is a website that contains real-world, self-reported salary data and is a major resource for economists studying pay inequality such as the gender pay gap. That is why we chose this dataset.

Exploratory Data Analysis: 1)Gender Counts 2)Sort for Gender Job Title, and BasePay 3)Add LogBasePay column and Gender Coded column [Male = 1, Female = 0]

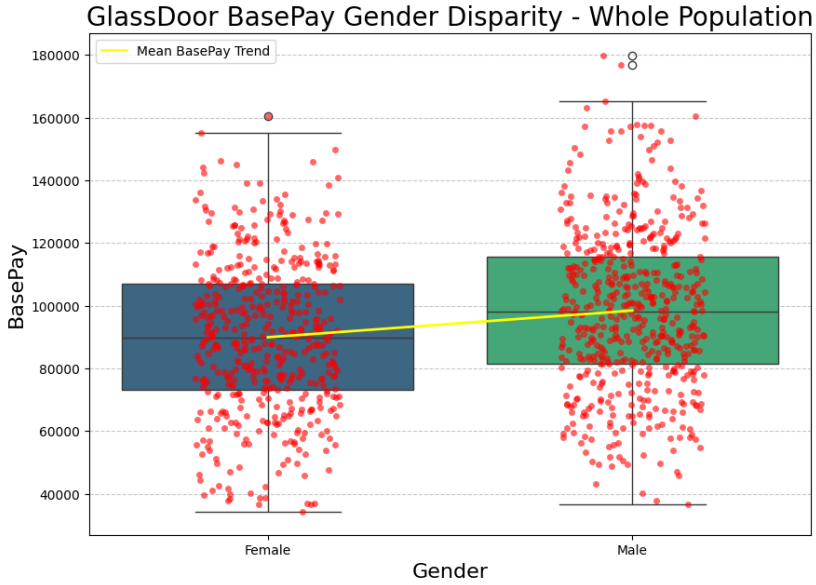


Figure 1: Pay Gap Across All Jobs

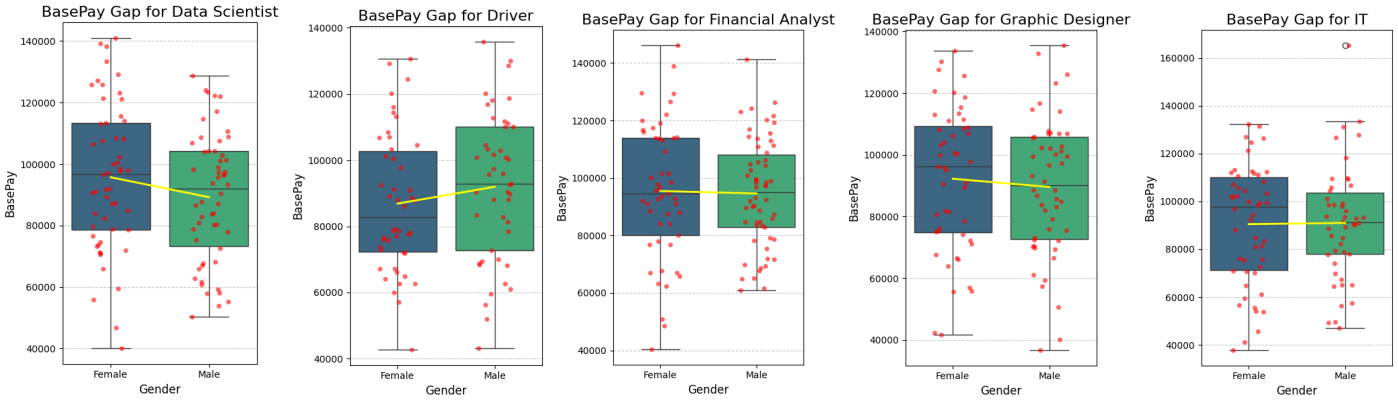
The average BasePay of Males and Females in the data set were calculated and connected by a line.

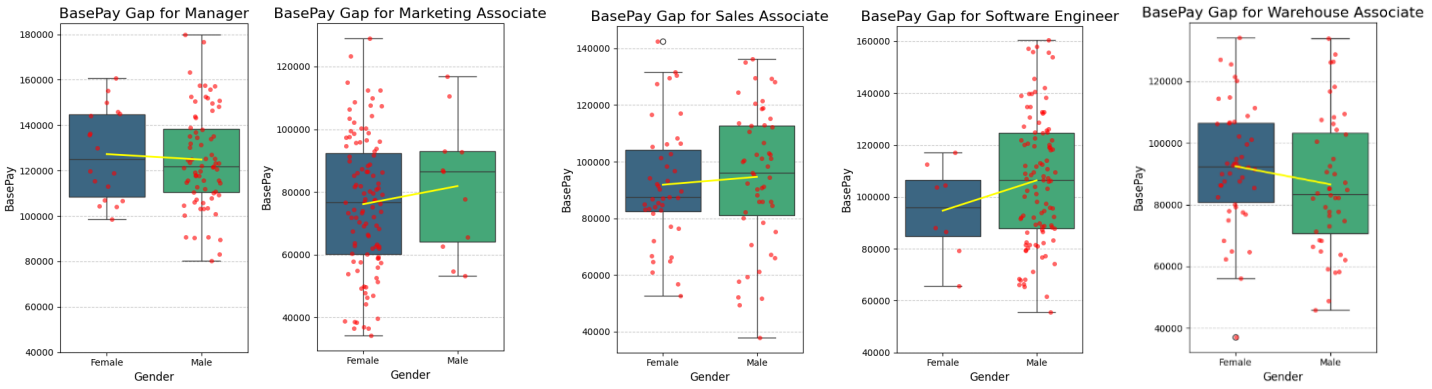
Males appear to earn around 10K more than females in 2019.

Figure 2-11: Job Title Pay Gaps

The average BasePay of Males and Females in each Job Title category of the data set were calculated and connected by lines.

The pay gap between genders appears to differ.





Methodology

(1) *Uncontrolled GLM for Whole Population* **[LogBasePay = b0 + b1*MALE]**

- **Intercept/b0** = 11.3665 (Females earn a LogBasePay of 11.37 on average) p>|t| = 0.000
- **MALE** = 0.0953 (9.53% Inc. in BasePay if Male) p>|t| = 0.000
- **Conclusion:** Males seem to earn more than women in general

(1a) *GLM for Job Title Specific Subsamples* **[LogBasePay = b0 + b1*MALE, data = Job Title]**

- Noticed some negative correlations in the MALE coefficients for particular jobs

(2) *Controlled GLM for Job Title* **[LogBasePay = b0 + b1*MALE + b2*(Jobtitle 1) + b3*(Jobtitle 2) …]**

- **Intercept/b0** = 11.4036 (Female Data Sci. earn LogBasePay of 11.40 on average) p>|t| = 0.000
- **MALE** = -0.0030 (-0.3% Dec. in BasePay if Male) p>|t| = 0.869
- **Conclusion:** Men DO NOT consistently earn more than women, but we FAIL TO REJECT our null hypothesis
- **Note:** This model uses Data Scientist as a reference and is included within the Intercept Value
- **Note:** The MALE coefficient therefore, calculates the % difference in job earnings vs. Female Data Scientists

Results & Analysis

After controlling for job title with the second GLM, the gender pay gap changed drastically. The MALE coefficient decreased in magnitude and became negative. This means that males actually earned less than females in particular jobs compared to their female data scientist counterparts, that being the Graphic Designer, Driver, IT, Marketing Associate and Warehouse Associate roles specifically.

The p-value for the MALE coefficient value did not meet the 95% confidence level (0.869), meaning that if the null hypothesis were true, there would be an 86.9% probability of observing an average Gender pay gap as extreme as, or more extreme than, the coefficient observed.

The residual plots for the all job titles were relatively homoscedastic, normalized, and independent–evenly distributed across the y=0 line with no obvious residual trend. This means that our measurements/predictions from the models are valid.

Conclusions

In this research project, we were curious whether the gender pay gap changed if we controlled for Job Title. In order to answer this question, we took salary data from Glassdoor up to 02/28/2019 and created boxplot-strip plot visualizations to represent correlations in the data. We found that males generally earned more than females, but once we controlled job title by running a sophisticated GLM, workers that were male actually earned less than their female

counterparts in particular jobs as shown by the MALE coefficient changing substantially from positive to negative. The residual plots confirmed that our models predictions were valid via linearity, normality, and homoskedasticity. While the p-value was high, forcing us to fail to reject the null hypothesis, what we can conclude Men DO NOT consistently earn more than women when controlling for Job Titles.

References

1) Aragão, R. F. and C. (2025, March 4). Gender pay gap in U.S. has narrowed slightly over 2 decades. *Pew Research Center*. https://www.pewresearch.org/short-reads/2025/03/04/gender-pay-gap-in-us-has-narrowed-slightly-over-2-decades/

2) Glassdoor. Research, (2025) https://www.glassdoor.com/blog/category/research/

3) Jauhari, Neelima. Glassdoor: Analyze Gender Pay Gap– Kaggle, (2019) https://www.kaggle.com/datasets/nilimajauhari/glassdoor-analyze-gender-pay-gap