Analysis of Potential Factors Affecting Salary

A Nested and Crossed ANOVA Study Based on Permanent Employees of Montgomery County, MD.

Hao Zhou, Zhengxiao Yang

April 29, 2024

BIOS 7080 Design of Experiments

1 Introduction

In recent years, the topic of equitable salary distribution has been increasingly discussed in many public sectors. Government agencies have become the focus of attention because they play a leading role in matters related to employment practices, including how much they pay their employees. Montgomery County government is an suitable case study candidate due to its hierarchical structure of domain, department and division[1]. Analyzing its salary distribution is important to understand how structural and demographic variables affect salaries.

The main goal of this study was to examine how key structural and demographic factors (including domain, department, division and gender) influenced the salary status of permanent employees in Montgomery County, using data from the 2023 government release. To ensure accuracy and robustness, data preprocessing was conducted to address missing data and potential outliers. This study adopted a Nested and Crossed Analysis of Variance (ANOVA) design[2] and constructed a sampling plan that reflected the inherent hierarchical nature of the data. This design allowed for a more nuanced exploration of the interactions between factors.

The results revealed patterns of salary disparities and helped propose recommendations for developing better strategies regarding salary equity in the public sector. Given growing calls for greater scrutiny and transparency, this analysis of factors influencing salary statuses provided valuable lessons that could be extended to other government entities to optimize their pay strategies.

2 Data Preprocessing

2.1 Data Cleaning

In the data cleaning phase of our analysis, apart from the usual procedures for handling missing and outlier values[3], we employed various other principles that best fitted the context of our organizational study. Thus, we had:

1. **Single-Gender Divisions**: Divisions comprising only male or female employees were excluded. Often, such homogeneity is due to particular contextual reasons which may be biased in conducting the analysis.

2. **Small Divisions**: Divisions with very small employee counts were excluded. In case of a small sample size, statistical distribution becomes unreliable due to small amounts of data, which affects the robustness.

3. **Small Departments**: We excluded departments that contained only a small number of divisions. Mostly, departments with minimal divisions tend to lack enough variability or adequate scope of data in meaningfully analyzing how structural factors of the department influence salary distributions.

The intention of these principles is to make the dataset robust enough to reflect the complexities and variations in the salary distributions within an organization, representative, and reliable for analysis.

2.2 Sampling

In this study, the employees are in a hierarchical structure. The sampling that has to be done to achieve the experimental design needs to be well-organized.

The sampling used both stratified and cluster sampling methods[4][5]. First, all the departments were stratified into 3 domains: Public Safety, Administrative Services, and Community Services. From each domain, 3 departments were randomly selected. Then, cluster sampling was used to obtain 6 different divisions in each department and 6 employees in each division. Significantly, stratified sampling ensured equal representation of male and female employees. More specifically, in each division, there was the sampling of 3 male and 3 female employees, respectively (see Figure 1).

This complex sampling scheme effectively reflects the organization's complexity and ensures a dataset capable of rigorously analyzing the interaction of structural and demographic factors on salaries.

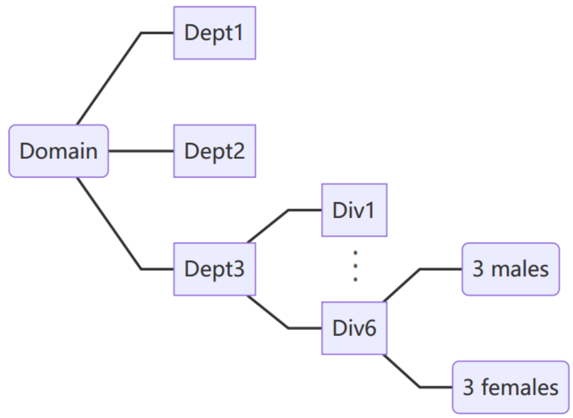

Figure 1: Sampling Process

Considering Figure 2, by comparing the distribution of total salary in the sample and in the population, we believe that the results of the sampling are a good representation of the original distribution and the sampling process can be considered appropriate.

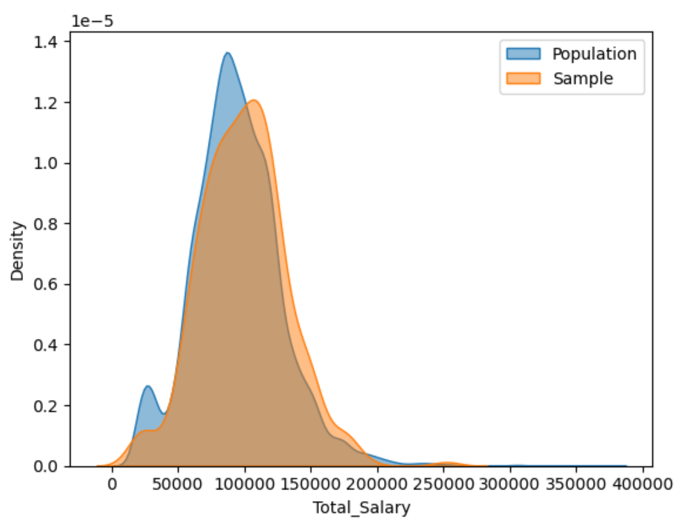

Figure 2: Distributions in Sample and Population

3 Model

3.1 Mathematical Model

The ANOVA model includes Gender (A), Domain (B), Department (C) nested within Domain, and Division (D) nested within Department. The effects are modeled as fixed for Gender and Domain, and random for Department and Division:

$$y_{ijk\ell n} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \gamma_{jk} + (\alpha\gamma)_{ijk} + \delta_{jk\ell} + (\alpha\delta)_{ijk\ell} + \varepsilon_{ijk\ell n},$$

$1 \leq i \leq 2, 1 \leq j \leq 3, 1 \leq k \leq 3, 1 \leq l \leq 6, 1 \leq n \leq 3.$

- $y_{ijk\ell n}$: The salary of a single employee.
- μ: The overall mean salary across all domains and genders.
- α_i: The fixed main effect of gender.
- β_j: The fixed main effect of domain.
- $(\alpha\beta)_{ij}$: The fixed interaction effect between gender and domain.
- γ_{jk}: The random effect of the kth department within the jth domain, which follows $\mathrm{N}(0, \sigma_\gamma^2)$.
- $(\alpha\gamma)_{ijk}$: The random interaction effect between gender and department within domain, which follows $\mathrm{N}(0, \sigma_{\alpha\gamma}^2)$.
- $\delta_{jk\ell}$: The random effect of the ℓth division within the kth department in the jth domain, which follows $\mathrm{N}(0, \sigma_\delta^2)$.
- $(\alpha\delta)_{ijk\ell}$: The random interaction effect between gender and division within department, which follows $\mathrm{N}(0, \sigma_{\alpha\delta}^2)$.
- $\varepsilon_{ijk\ell n}$: The random error term, which follows $\mathrm{N}(0, \sigma^2)$.

3.2 Assumptions

1. **Lack of fit**: It is necessary to check whether the model fits well. The random pattern of standardized residuals by gender and division (Figure 3.a and Figure 3.b) does not indicate any significant lack of fit.

2. **Independence**: Observations within and across cells must be independent. It could be checked by plotting residuals against the order of observations in the original dataset (Figure 3.c). The random pattern of standardized residuals suggests that the observations are independent of each other.

3. **Normality**: The standardized residuals of the model should follow a normal distribution. The Q-Q plot (Figure 3.d) shows that this assumption is satisfied.

4. **Homogeneity of Variances**: Observations in all cells must have constant variances, which is a crucial assumption for the pooled variance used in ANOVA. Levene's test (Med) is used to test this assumption. The result is: test statistic = 0.6876, p-value = 0.9849 > 0.05. Therefore, the variances of different cells should be considered constant.

3.3 EMS and Test Statistics

Table 1 presents the expected mean squares (EMS) and corresponding test statistics for each source of variability. Note that $\theta_\alpha^2 = \sum_{i=1}^2 \alpha_i^2$ and $\theta_\beta^2 = \sum_{j=1}^3 \beta_j^2/2$.

Table 1: EMS and Test Statistics

Source	**EMS**	**Test Statistic**
Gender	$\sigma^2 + 3\sigma_{\alpha\delta}^2 + 18\sigma_{\alpha\gamma}^2 + 162\theta_\alpha^2$	$\text{MS}_A/\text{MS}_{AC(B)}$
Domain	$\sigma^2 + 3\sigma_{\alpha\delta}^2 + 18\sigma_{\alpha\gamma}^2 + 6\sigma_\delta^2 + 36\sigma_\gamma^2 + 108\theta_\beta^2$	$\text{MS}_B/\text{MS}_{C(B)}$
Gender × Domain	$\sigma^2 + 3\sigma_{\alpha\delta}^2 + 18\sigma_{\alpha\gamma}^2 + 54\theta_{\alpha\beta}^2$	$\text{MS}_{AB}/\text{MS}_{AC(B)}$
Dept (Domain)	$\sigma^2 + 3\sigma_{\alpha\delta}^2 + 18\sigma_{\alpha\gamma}^2 + 6\sigma_\delta^2 + 36\sigma_\gamma^2$	$(\text{MS}_{C(B)} + \text{MS}_{AD(C)})/(\text{MS}_{AC(B)} + \text{MS}_{D(C)})$
Gender × Dept (Domain)	$\sigma^2 + 3\sigma_{\alpha\delta}^2 + 18\sigma_{\alpha\gamma}^2$	$\text{MS}_{AC(B)}/\text{MS}_{AD(C)}$
Div (Dept)	$\sigma^2 + 3\sigma_{\alpha\delta}^2 + 6\sigma_\delta^2$	$\text{MS}_{D(C)}/\text{MS}_{AD(C)}$
Gender × Div (Dept)	$\sigma^2 + 3\sigma_{\alpha\delta}^2$	$\text{MS}_{AD(C)}/\text{MS}_E$
Residual	σ^2	

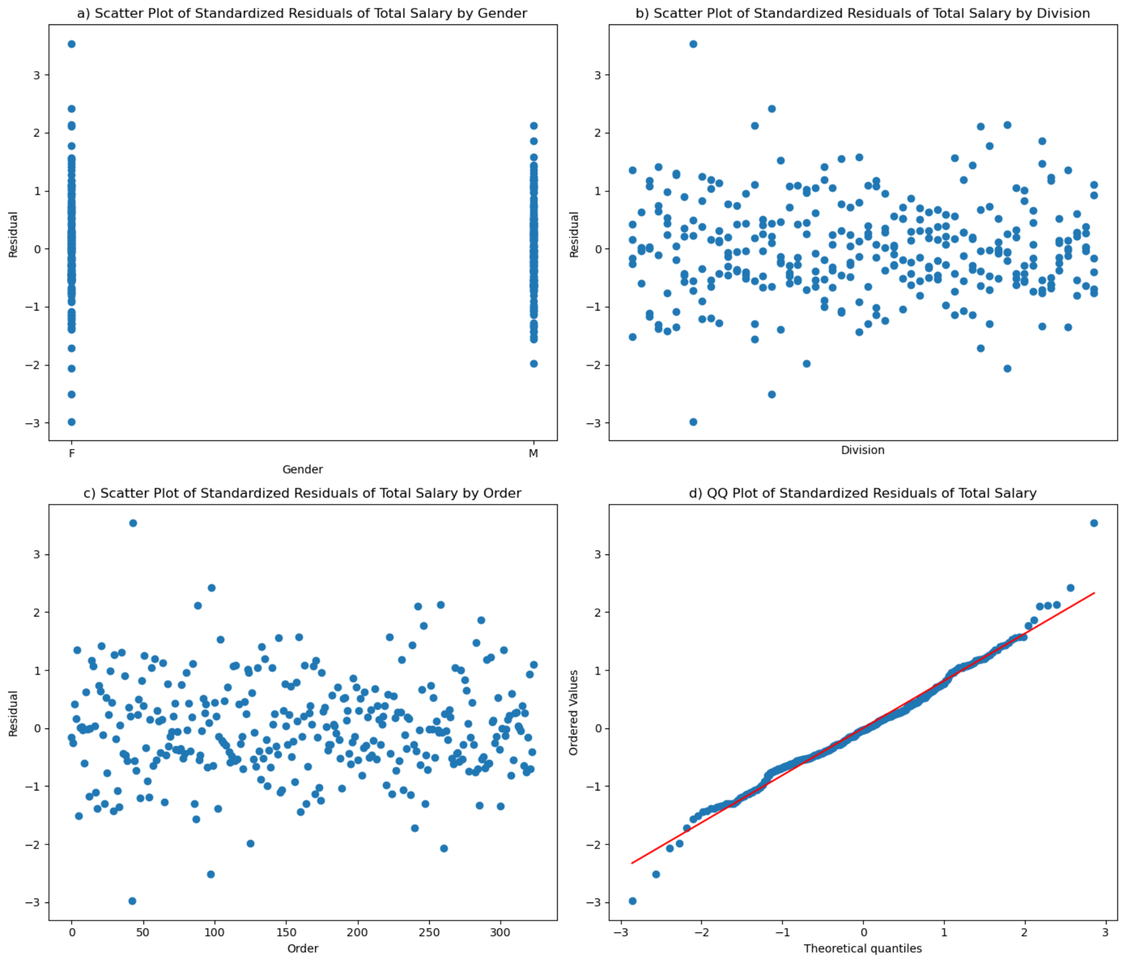

Figure 3: Residual Diagnostic Plots

4 Results

4.1 ANOVA Table

Python was used for programming[6], with several key libraries imported[7–10]. The results are presented in Table 2.

Table 2: ANOVA Summary Table

Source	**df**	**SS**	**MS**	**F**	**PR(>F)**
Gender	1	4.558272e+08	4.558272e+08	0.619350	0.461234
Domain	2	2.057873e+10	1.028937e+10	1.184006	0.368627
Domain × Gender	2	2.230723e+09	1.115361e+09	1.515485	0.293258
Dept (Domain)	6	5.214181e+10	8.690302e+09	3.865940	0.002999**
Gender × Dept (Domain)	6	4.415859e+09	7.359766e+08	1.200868	0.323394
Div (Dept)	45	7.517109e+10	1.670469e+09	2.725647	0.000522***
Gender × Div (Dept)	45	2.757917e+10	6.128705e+08	0.822123	0.780640
Residual	216	1.610222e+11	7.454733e+08		

4.2 Post-hoc Analysis

After obtaining significant information, the model was reduced to: $y_{ijkln} = \mu + \gamma_{jk} + \delta_{jkl} + \varepsilon_{ijkln}$. Then standard deviations from different sources could be estimated[11]: $\hat{\sigma}_\gamma = 13841, \hat{\sigma}_\delta = 13277, \hat{\sigma} = 27303$. Thus, 95% confidence intervals of average salaries of different departments could be estimated[2] (see Table 3).

Table 3: CI for Average Salaries of Different Departments

Department Name	Lower Bound($)	Upper Bound($)
Alcohol Beverage Services	58,273	81,153
Correction and Rehabilitation	93,926	116,806
Department of Environmental Protection	93,958	116,838
Department of Finance	102,990	125,870
Department of General Services	89,326	112,206
Department of Permitting Services	104,872	127,752
Department of Police	90,656	113,536
Department of Recreation	64,972	87,853
Department of Transportation	92,163	115,043

we can calculate the power as well. The power of detecting a 20% increase in the overall standard deviation caused by domain at the 0.05 significance level, which could be stated as $\frac{\sqrt{\theta_\beta^2+\sigma_\gamma^2+\sigma_\delta^2+\sigma^2}}{\sqrt{\sigma_\gamma^2+\sigma_\delta^2+\sigma^2}} \geq 1.2$, is 0.52.

5 Discussion

5.1 ANOVA

The results of the Nested and Crossed ANOVA, as depicted in Table 2, illuminates the structural and demographic factors affecting salary distributions.

The significant main effect of the department nested within domain (denoted by the 'Domain:Department', p <0.01) points to the influence of departmental classification on salary differences. This indicates that, within each domain, the department where an employee is positioned plays a substantial role in determining salary.

Equally telling is the significant main effect of the division nested within department (denoted by 'Domain:Department:Division', p <0.001). The highly significant p-value suggests that within each departments, the divisional breakdown yields substantial variations in salaries. This may reflect the internal hierarchy and the specific roles that are unique to each division within the government structure.

Examining the role of gender and domain, our analysis reveals no significant main effect or interaction effects with other factors. This finding may be interpreted as an indication that Montgomery County's government does not exhibit salary disparities solely based on domain and gender. This seems to contradict the expectations that arise from discussions in contemporary society, particularly about the important role of gender equality in the issue of remuneration[12]. However, it is essential to approach this result with a nuanced understanding that the absence of a statistically significant difference does not unequivocally prove the absence of all gender-related salary disparities, but rather that such disparities, if present, are not manifesting at a detectable level within the scope of the factors considered in this model.

Without considering exogenous biases, one possible explanation is that we must take into account the special characteristics of the government sector. That is, unlike other organizations with a similar structure, it must first and foremost consider the issue of equity. Differences in salary due to position and personal characteristics are more legitimate than differences in domain and gender, and do not violate the social emphasis on fairness. In terms of results, they are closer to a natural state of affairs, which shows significance.

5.2 Post-hoc Analysis

The post-hoc analysis serves as a natural extension of the ANOVA[11], further dissecting the sources of salary variation within Montgomery County's government. The standard deviation estimates reflect the variability of salaries within departments and divisions. The relative sizes of $\hat{\sigma}_\gamma$ and $\hat{\sigma}_\delta$ compared to $\hat{\sigma}$ suggest that while both factors contribute significantly to the variance, there's still a large portion of variance unexplained by these factors. It might be beneficial to investigate other potential variables, interactions, or higher-level effects not currently included in the model.

The gap between departments is demonstrated in Table 3. Some departments, such as Alcohol Beverage Services and Department of Recreation, have CI upper bounds failing to reach the lower bounds of others, exhibiting significantly lower salaries. The source of this may partially be explained by the components of salaries. Total salaries consist of three parts: the basic salary, overtime pay, and longevity pay. Where longevity pay are potentially related to the start-up of the department, active years, and changes in employee demand. And overtime pay is potentially related to various types of unforeseen events.

While a significant effect of domain is not detected, it does not necessarily mean that the effect is truly non-significant. This may be due to the insufficient sample size and the low power (0.52) resulting from it. So the sample size should be adjusted to achieve a higher power (typically 0.8) in further study. Two possible sampling schemes of a power of 0.8 are given:

(1) 8 departments from each domain, 4 divisions from each department, 6 employees from each division;

(2) 7 departments from each domain, 7 divisions from each department, 7 employees from

each division.

Note that gender is no longer included as a factor in the sampling structure, as it proved to be insignificant.

5.3 Limitations

Data Quality: Data is only available from the Montgomery County government for the year 2023, which means that extrapolation of the results to more general areas needs to be approached cautiously.

Balanced Sampling: In our design of experiment, a sampling method combining cluster sampling and stratified sampling is proposed. From the principle, the sample size taken should be weighted. However, for simplicity purposes, a balanced sample was simply used in the sampling and no corrections were made to the results. Although a balanced sample is a good approximation in some cases, the influence on the results is difficult to avoid.

Robustness: ANOVA is primarily used to detect differences in group means and is not suitable for other types of differences, such as medians. However, as a measure of central tendency, the mean has a breakdown value of 0 and is not robust. This is particularly evident in data with high variability, which can lead to incorrect interpretations of differences between groups.

5.4 Future Directions

Expand Data Sources: To increase the generalizability of the study, future studies should consider expanding the data sources. Including data from multiple regions or multiple time points.

Implement Weighted Sampling Techniques: Future research should use weighted sampling techniques to more accurately represent the population structure. This includes adjusting the sample sizes for each stratum or cluster so that they better reflect their proportion in the overall population.

Utilize Robust Statistical Methods: Robust statistical methods could be employed. For example, rank-based non-parametric methods, such as the Kruskal-Wallis test[13]. Additionally, exploring data transformation strategies (such as log transformation) to stabilize variance and reduce the impact of skewness might be an effective way to enhance the accuracy of analyses.

6 Conclusions

Key results from this Nested and Crossed ANOVA on salary data of permanent employees in Montgomery County are that the structures of department and division make a substantial contribution to the variation in salary, whereas the effects of gender and domain are minor. Specifically, salaries in Alcohol Beverage Services and Department of Recreation are significantly lower compared to other departments.

Future studies should expand their sources of data to a wider range of time points and geographical regions for enhancing generalizability and validity. More factors, such as educational background, work experience, and other demographic variables, could be included to provide a better insight into the determinants of salary inequality in the public sectors. This would refine the existing models and support the development of more comprehensive strategies for managing salary disparities in governmental entities.

References

[1] Montgomery County Government. *Employee Salaries - 2023*. `https://catalog.data.gov/dataset/employee-salaries-2023`. Accessed: Apr 2024. 2023.

[2] Mark Anderson and Patrick Whitcomb. "Design of Experiments: Statistical Principles of Research Design and Analysis:Design of Experiments: Statistical Principles of Research Design and Analysis". In: *Technometrics* 43 (May 2001), pp. 236–237. DOI: `10.1198/tech.2001.s589`.

[3] Hadley Wickham. "Tidy data". In: *Journal of statistical software* 59 (2014), pp. 1–23.

[4] Paul S Levy and Stanley Lemeshow. *Sampling of populations: methods and applications*. John Wiley & Sons, 2013.

[5] Sharon L Lohr. *Sampling: design and analysis*. Chapman and Hall/CRC, 2021.

[6] *Python Programming Language*. `https://www.python.org`.

[7] Charles R. Harris, K. Jarrod Millman, Stéfan J. van der Walt, et al. *Array programming with NumPy*. 2020. URL: `https://doi.org/10.1038/s41586-020-2649-2`.

[8] Wes McKinney. "Data Structures for Statistical Computing in Python". In: *Proceedings of the 9th Python in Science Conference*. Ed. by Stéfan van der Walt and Jarrod Millman. 2010, pp. 51–56. URL: `https://conference.scipy.org/proceedings/scipy2010/pdfs/mckinney.pdf`.

[9] John D. Hunter. "Matplotlib: A 2D Graphics Environment". In: *Computing in Science Engineering* 9.3 (2007), pp. 90–95. URL: `https://doi.org/10.1109/MCSE.2007.55`.

[10] Skipper Seabold and Josef Perktold. *Statsmodels: Econometric and Statistical Modeling with Python*. 2010. URL: `https://conference.scipy.org/proceedings/scipy2010/seabold.html`.

[11] Douglas C Montgomery. *Design and analysis of experiments*. John wiley & sons, 2017.

[12] Francine D. Blau and Lawrence M. Kahn. "The Gender Wage Gap: Extent, Trends, and Explanations". In: *Journal of Economic Literature* 55.3 (Sept. 2017), pp. 789–865. DOI: `10.1257/jel.20160995`. URL: `https://www.aeaweb.org/articles?id=10.1257/jel.20160995`.

[13] William H Kruskal and W Allen Wallis. "Use of ranks in one-criterion variance analysis". In: *Journal of the American statistical Association* 47.260 (1952), pp. 583–621.