# **Chapter 4**

Gina Marshall-Johnson

Omega Graduate School

Joshua Reichard, PhD

August 27, 2024

## Introduction

The purpose of this study is to examine the association in interview selection rates among marginalized groups in STEM fields according to the use of Generative AI for resume creation or enhancement. This chapter covers the data collection, preparation and analysis of this study. There were seven hypotheses in this study. There was one hypothesis which examined whether a statistically significant association existed in interview preselection with the use of Generative AI. There were two hypotheses which examined whether there were statistically significant differences between interview preselection and the use of Generative AI with particular focus on the race of historically marginalized populations. There were two hypotheses which examined whether there were statistically significant associations between interview preselection and the use of Generative AI with particular focus on the gender of historically marginalized populations. There was one final hypothesis which examined the perceived accuracy of Generative AI to the user, independent of interview preselection.

The study examined thirty-five responses in an on-line text-based survey. The data were collected over a five-week period. Of the respondents 66% were from historically marginalized populations. These provided a strong foundation on which to compare the impact of Generative AI on both interview preselection and perceived accuracy. 43% of the respondents were female. Slightly more than half or 57% used Generative AI and correspondingly 43% did not use Generative AI in preparation of their resume. There was a nearly equal distribution from which to compare the impact on the use of Generative AI. 71% of respondent were selected for an interview which was reflective of the survey being distributed primarily to those who were recent hires within STEM companies. This may reflect sample bias. 83% of historically marginalized samples who used Generative AI were selected for interviews. This was likely a reflection of the sample being recently employed in STEM. Interestingly, a larger percentage of whites (45%) and males (81%) did not use Generative AI and were selected for interviews. There were an equal percentage of females selected for interviews as males even though only 18% of females did not use Generative AI.

## Deviations in Data Collection and Preparation

Data was downloaded into a CSV file and cleaned. Of the 37 responses, there were two that were eliminated because they were generated during system and process testing. Thus, the sample analyzed consisted of 35 survey respondents. Of the remaining 35 responses one had a field left blank on gender. There was an option for gender “prefer not to say” so the blank field was assumed to be this option. On question five, Generative AI was accurate in describing my qualifications, two respondents skipped this question. Those that skipped had not used Generative AI. The analysis was performed on the results received. Race was identified in 100% of the responses. Twenty-nine of the thirty-five respondents applied for a job in STEM. Twenty of the respondents used Generative AI to assist in the preparation of their resume. The professional snowball sampling yielded over 50% of responses from historically marginalized samples, which was greater than the average working in STEM. The anomaly provides a strong basis from which to analyze the impact of Generative AI’s use on historically marginalized samples. Snowball sampling was likely centered in the computer science field due to the immediate social network of the researcher.

## Results

The results of this quantitative, quasi-experimental design which examined interview selection rates for statistically significant associations among diverse job seekers in STEM fields who utilize Generative AI for resume creation or enhancement and those who do not are discussed in this section. The data collected was from college-educated STEM job seekers. This section discusses survey sample demographics, descriptive statistics and hypothesis testing.

**Assumption Tests**

Of the eight survey questions, seven were categorical responses. The data from the non-categorical response was tested for normality. This was a Likert question that asked the degree to which Generative AI helped in their application process, if they used Generative AI. The type of test used to test normality was a visual inspection of a histogram. The data were not normally distributed based on visual inspection of this histogram analysis. Non-parametric means were used to analyze the data consequently. While collecting this data were informative, it was not used to answer any of the research questions.

**Participant Demographics**

The survey demographics were plotted based on quantity of responses and percentages are found in the following figures and tables. Table 1 contains the respondent percentages who applied for a job in STEM within the past twelve months.

**Table 1**

*Survey Respondents Distribution by STEM Position Application*

|  |  |
| --- | --- |
| Respondents STEM Position Application | Percentage (%) |
| Applied for a job in STEM | 83 |
| Did not apply for a job in STEM | 17 |

Figure 1 depicts graphically 29 of the 35 respondents applied for a job in STEM within the past 12 months.

**Figure 1**

*Survey Respondents by STEM Position Application*

The racial demographics of respondents are shown in Table 2. The racial profile of responses was 40% for Black, 34% for White, 6% for Hispanic, and 11% for Asian responses. There were 6% from two or more races, 3% from other, and 0% from the prefer not to say category. The literature review found 9% Blacks working in STEM and in computer-related fields 7%. Data collection, using Snowball sampling produced a significantly higher percentage of Black responses. It was more than four times higher than the industry average. Similarly, the number of Hispanics working in STEM is 8%, and in computer-related fields is 8%. The responses were 6% for Hispanics which was closer to the industry average but slightly lower. The number of Hispanics working at Facebook, Google, and Microsoft are 5%, 3.6% and 6%, respectively (West et.al., 2019, p. 10). The survey demographics for Hispanic respondents are the same as workers reported at Microsoft. The literature review found the percentage of Asians working in STEM is 10% compared to 6% in all other fields. The survey responses were 11%, which was very close to industry averages. The only diverse racial demographic that was different from the industry average was the percentage of Black respondents. The study was centrally focused on the impact of Generative AI on historically underserved samples applying for jobs in STEM. Therefore, the data collected was particularly useful for this study. It may also be hard to duplicate.

**Table 2**

*Survey Respondents Racial Demographic Distribution*

|  |  |
| --- | --- |
| Respondents STEM Position Application | Percentage (%) |
| White/Caucasian | 34 |
| Black/African American | 40 |
| Hispanic/Latino | 6 |
| Asian/Pacific Islander | 11 |
| 2 or More | 6 |
| Other | 3 |
| Prefer not to say | 0 |

**Figure 2**

*Survey Respondents by Racial Distribution*

Table 3 depicts the gender demographics of the survey respondents.

**Table 3**

*Survey Respondent's Gender Distribution*

|  |  |
| --- | --- |
| Respondents Gender Distribution | Percentage (%) |
| Male | 57 |
| Female | 43 |

The survey responses were 57% from males and 43% from females. Literature shows that the number of women working in Computer Science has reduced from 32% down to 25% between 1990 and 2019. Computer Science has exploded in the number of jobs over this period. Having 43% responses is also higher than the 25% working in computer science fields, but similar to the life and physical sciences (48% and 40%, respectively). The percentage of survey responses for women was also like the 47% working in math fields (Kennedy, 2021, p. 1).

Figure 3 is a graphical representation of the number of men and women responding to the survey. While there were questions on non-binary and prefer not to say. This selection was not chosen by any of the respondents.

**Figure 3**

*Survey Respondents by Gender*

The demographics of race and gender are representative of the target sample for the study.

Table 4 depicts the data distribution in percentage for those who used Generative AI to write, create, or revise their resume.

**Table 4**

*Survey Respondents Use of Generative AI Distribution*

|  |  |
| --- | --- |
| Gen AI Use  | Percentage |
| Used | 57.1 |
| Did not use | 42.9 |

There is 57.1% of the respondents used Generative AI for their resume creation, revision or enhancement. There were 42.9% of respondents who did not use Generative AI for their resume creation, revision or enhancement. Given Generative AI was released to the public through ChatGPT in November of 2022, just 19 months before the survey, and it was used by 57% of the respondents indicates the target sample in STEM are early adopters and provides a representative sample for the impact of its use.

**Figure 4**

*Survey Respondent's use of Generative AI in the Application Process*



Non-Gen AI Use

Gen AI Use

The distribution for the interview selection for survey respondents is shown in Table 5.

**Table 5**

*Survey Respondents Interview Selection Distribution*

|  |  |
| --- | --- |
| Interview Selection | Percentage |
| Selected | 71.4 |
| Not Selected | 28.6 |

**Figure 5**

*Survey Respondents Interview Section Distribution*



Not selected for an interview

Selected for an Interview

There is 71.4% of the respondents were selected for an interview. There were 28.6% of respondents who were not selected for an interview.

The sample shown in Table 6-9, and Figure 6-9 further subdivides the total sample by those who applied for a job in STEM in the past 12 months n=29), used Generative AI in document preparation n=14) and were selected for an interview n=12).

Table 6 displays the racial demographic distribution for those who applied for a job in STEM in the past 12 months, used Generative AI to write, create or revise their resume and were selected for an interview. The final sample size was n=12.

**Table 6**

*Survey Respondents Racial Distribution for Interview Selection and Use of Generative AI*

|  |  |
| --- | --- |
| Racial Distribution | Percentage |
| White | 17% |
| Black | 42% |
| Hispanic | 0% |
| Asian | 25% |
| 2 or More Races | 8% |
| Other | 8% |
| Prefer not to Say | 0% |

**Figure 6**

*Survey Respondents Racial Distribution for Interview with the use of Gen AI*



The demographics for the sample that used Generative AI and were selected for an interview, 17% were White, 42% were Black, 0% were Hispanic, 25% were Asian, 8% were two or more races, 8% were other, and none selected prefer not to say. There were no Hispanics who applied for a job in STEM. There were two who used Generative AI in employment document preparation, however.

Table 7 displays the racial demographic distribution for those who did not use Generative AI to write, create or revise their resume and were selected for an interview n=13). The sample selected for an interview was n=11.

**Table 7**

*Survey Respondents Racial Distribution for Interview Selection did not use Generative AI*

|  |  |
| --- | --- |
| Racial Distribution | Percentage |
| White | 45.4% |
| Black | 36.4% |
| Hispanic | 0.00% |
| Asian | 0.00% |
| 2 or More | 9.1% |
| Other | 9.1% |
| Prefer not to Say | 0.00% |

**Figure 7**

*Survey Respondents Racial Distribution for Interview - did not use Gen AI*



The sample which applied for a job in STEM in the past 12 months, did not use Generative AI in their employment document preparation and were selected for an interview was 45.5% were White, 36.3% were Black, 9.1% were Hispanic, 9.1% were two or more races, 0% were other or prefer not to say. There were no Asians, who did not use Generative AI in the preparation of their resume and were selected for an interview in STEM.

Table 8 displays the gender distribution for those who did not use Generative AI to write, create or revise their resume and were selected for an interview. The sample analyzed was N=11.

**Table 8**

*Survey Respondent's Gender Distribution for Interview Selection - did not use Generative AI*

|  |  |
| --- | --- |
| Gender Distribution | Percentage |
| Male | 81.8% |
| Female | 18.2% |

**Figure 8**

*Survey Respondent's Gender Distribution for Interview - did not use Gen AI*



The demographics of those who applied for a job in STEM in the past 12 months, and did not use Generative AI in the preparation of their employment documents were 73% male and 27% female. The sample size was n=11. There were far fewer women than men who did not use Generative AI in the preparation of their resumes and were selected for an interview. This may be reflective of the lower number of women in the STEM field.

Table 9 displays the gender distribution for those who applied for a job in STEM in the past 12 months, used Generative AI to write, create or revise their resume and were selected for an interview. The final sample was n=12.

**Table 9**

*Survey Respondents Gender Distribution for Interview Selection with the use of Generative AI*

|  |  |
| --- | --- |
| Gender Distribution | Percentage (%) |
| Male | 50% |
| Female | 50% |

**Figure 9**

*Survey Respondent's Gender Distribution for Interview with Use of Gen AI*



There were an equal number of women and men who used Generative AI in the preparation of their resume and were selected for an interview.

The sample size for those who used Generative AI and responded regarding the resulting accuracy of the model was n=13. The racial demographic percentage summary for respondents’ perception of accuracy is in Table 10. The racial percentage distribution was 23% White, 38% Black, 8% Hispanic, 23% Asian, and 8% two or more races consider Generative AI’s description of their qualifications accurate. Figure 10 graphically displays the racial demographic distribution.

**Table 10**

*Survey Respondents' Racial Distribution Perceived Accuracy of Generative AI*

|  |  |
| --- | --- |
| Racial Distribution | Percentage |
| White | 23% |
| Black | 38% |
| Hispanic | 8% |
| Asian | 23% |
| 2 or More | 8% |
| Other | 0% |
| Prefer not to Say | 0% |

**Figure 10**

*Survey Respondents Racial Distribution for Perceived Accuracy with the use of Gen AI*



**Table 11**

*Survey Respondents Gender Distribution on Perceived Accuracy of Generative AI*

|  |  |
| --- | --- |
| Gender Distribution | Percentage |
| Male | 38% |
| Female | 62% |

**Figure 11**

*Survey Respondents' Gender Distribution Perceived Accuracy of Gen AI*



**Descriptive Statistics**

The analysis of descriptive statistics discussion was limited because all the data collected was binary, except for one question. All the research questions being considered are based on categorical, binary data.

There was one Likert question for which a statistical distribution can be measured. The question assesses the degree of helpfulness of Generative AI in preparing pre-employment documentation. Using a histogram for visual analysis, it does not appear to be a normal distribution. A figure of the distribution is in Figure 12.

**Figure 12**

*Survey Respondents' Gender Distribution Perceived Accuracy of Gen AI*



## Findings

The seven hypotheses statistical analysis was performed using Chi-Squared, ANOVA or Mann-Whitney U analyses to determine whether there was a significant association between categorical variables. While Chi-Squared does not require equal sample sizes it does require enough observations in each category. The observation focused on whether the observed frequencies differ substantially from the expected frequencies.

Statistical procedures were conducted using PSSR (Practical Statistics for Social Research), statistical analysis software developed by Joshua D. Reichard for Omega Graduate School based on the jStat library (Reichard, 2024).

**Hypothesis 1**

**H01:** No statistically significant association exists between interview selection rates of those who use Generative AI for resume creation or enhancement and those who do not among job seekers in STEM fields (RQ1).

**Ha1:** A statistically significant association exists between interview selection rates of those who use Generative AI for resume creation or enhancement and those who do not among job seekers in STEM fields (RQ1).

**Table 12**

*2x2 Chi-Squared Matrix for Hypothesis 1*

|  |  |  |
| --- | --- | --- |
|  | Selected for Interview | Not Selected for Interview |
| Use of AI | 12 | 4 |
| No use of AI | 11 | 2 |
| Total | 23 | 6 |

***Hypothesis 1 Chi-Squared Contingency***

The Chi-Squared Contingency for Hypothesis One is shown in Table 13.

**Table 13**

*Chi-Squared Contingency Table*

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Selected | Not Selected |
| Group 1 | Used Gen AI | 12 | 4 |
| Group 2 | No Gen AI use | 11 | 2 |
|  | Total | 23 | 6 |

A Chi-Squared test of independence was conducted to examine the association between two categorical variables by comparing observed frequencies in a contingency table with expected frequencies derived under the hypothesis that the variables are independent. (MacFarland et al., 2016, p. 80; Creswell et al., 2023, p. 268).

**Table 14**

*Chi-Squared Summary Table*

|  |  |
| --- | --- |
| Statistic | Value |
| Chi-Squared Statistic | 0.4041 |
| Degrees of Freedom | 1.0000 |
| p-value | 0.52496666 |
| Fisher's Exact p-Value | 0.66283525 |

**Hypothesis 2**

**H02:** No statistically significant difference exists between interview selection rates and racial groups among candidates in STEM fields who use Generative AI for resume creation or enhancement. (RQ2).

**Ha2:** A statistically significant difference exists between interview selection rates and racial groups among candidates in STEM fields who use Generative AI for resume creation or enhancement (RQ2).

The sample size analyzed was n=16 with selected and not selected in same analysis by race. An ANOVA (Analysis of Variance) procedure was applied to the sample to examine whether a statistically significant difference exists between the means of 5 unrelated groups. (Creswell et al., 2023, p. 268).

***Racial Groups Examined (Independent Variable)***

The dependent variable of interview selection rates and the independent variable of racial group means are depicted in the Dependent Variable Summary by Independent Variable Groups in Table 15.

**Table 15**

*Dependent Variable Summary by Independent Variable Groups*

|  |  |  |
| --- | --- | --- |
| Group | n | Mean(x̄) |
| White | 5 | 0.4000 |
| Black | 6 | 0.8333 |
| Asian | 3 | 1.0000 |
| 2 or More | 1 | 1.0000 |
| Other | 1 | 1.0000 |

***ANOVA Statistics***

The ANOVA statistics are found in Table 16. There was insufficient evidence to reject the null hypothesis. This suggests that the variation within the groups was comparable to the variation between groups implying that any differences observed could likely be attributed to random chance rather than a true effect. Consequentially there was insufficient evidence to conclude that there was a statistically significant difference between the racial groups based on the use of Generative AI and interview selection.

**Table 16**

*ANOVA Results Statistics*

|  |  |
| --- | --- |
| Statistic | Value |
| F-Statistic | 1.307377 |
| p-Value | 0.32644373 |
| Omega Squared (ω²) | 0.071361 |
| LSD (post hoc) | 0.32644373 |

***Pairwise Comparisons for Significant Differences***

Pairwise comparisons in ANOVA were used to identify which specific groups differ significantly from each other by comparing the means of each pair of groups to determine where the significant differences lie. (Creswell et al., 2023, p. 268).

**Table 17**

*Pairwise Comparisons for Significant Differences*

|  |  |
| --- | --- |
| Group A | Group B |
| White | Asian |
| White | Two or More |
| White | Other |

There was a significant difference between the White group and the Asian, Two or more and Other (South Asian) groups. The significant differences in means could be attributed to a significantly higher percentage of whites who use Generative AI and were selected for interviews (45.4%) than Asian (0%), Two or More (9.1%) and Other (9.1%) who were selected for interviews.

***Between Groups Effects***

In the ANOVA analysis shown in Tables 18 and 19, the sum of the squares between racial groups (SSB) represents the variation due to the difference between the group means. It reflects how much time group means differ from the overall mean. The sum of the square within groups (SSW) also known at the sum of the squares for error, represents the variation within each group, indicating the variability among individual data points within the same group. The means square between groups (MSB) was calculated by dividing the SSB by the corresponding degrees of freedom, and the mean square within groups (MMW) was obtained by dividing the SSW by its degrees of freedom. These mean square values are used to calculate the F-statistic, which determines whether the variability between group means was significantly greater than the variability within groups, there by testing the null hypothesis. (Creswell et al., 2023, p. 268).

**Table 18**

*ANOVA Between Groups Effects*

|  |  |
| --- | --- |
| Measure | Value |
| Sum of Squares | 0.966667 |
| Mean Square | 0.241667 |
| Degrees of Freedom | 4.000000 |

The analysis of variance (ANOVA) was conducted to examine the differences between groups on the specified measure. The results are presented in Table 18 below, detailing the sum of squares, mean square, and degrees of freedom for the between-groups effects.

**Table 19**

*ANOVA within Groups Effects*

|  |  |
| --- | --- |
| Measure | Value |
| Sum of Squares | 2.033333 |
| Mean Square | 0.184848 |
| Degrees of Freedom | 11.000000 |

The results of the ANOVA statistical analysis were there was insufficient statistical evidence to reject the null hypothesis. The significant differences between group means were likely due to the small sample of some groups compared to others.

**Hypothesis 3**

**H03:** No statistically significant differences exist between interview selection rates and racial groups among candidates in STEM fields who do not use Generative AI for resume creation or enhancement (RQ3).

**Ha3:** A statistically significant difference exists between interview selection rates and racial groups among candidates in STEM fields who do not use Generative AI for resume creation or enhancement (RQ3).

The data set for hypothesis three was those who applied for a job in STEM within the past 12 months, did not use Generative AI for the creation, revision or enhancement of their resume and were selected for an interview. The sample size was n=13.

An ANOVA (Analysis of Variance) procedure was applied to the sample to examine whether a statistically significant difference exists between the means of four unrelated groups. (Creswell et al., 2023, p. 268).

***Groups Examined (Independent Variable)***

The dependent variable of interview selection rates and the independent variable of racial groups is depicted in the Dependent Variable Summary by Independent Variable Groups in Table 20.

**Table 20**

*Dependent Variable Summary by Independent Variable*

|  |  |
| --- | --- |
| Group | n |
| White | 6 |
| Black | 5 |
| Asian | 1 |
| 2 or More | 1 |

***ANOVA Statistics***

The ANOVA statistics are found in Table 21. There was insufficient evidence to reject the null hypothesis. This suggests that the variation within the groups was comparable to the variation between groups implying that any differences observed could likely be attributed to random chance rather than a true effect. Consequentially there is insufficient evidence to conclude that there was a statistically significant difference between the racial groups based on not using Generative AI and interview selection. (Creswell et al., 2023, p. 268).

**Table 21**

*ANOVA Results Statistics*

|  |  |
| --- | --- |
| Statistic | Value |
| F-Statistic | 0.108320 |
| p-Value | 0.95309464 |
| Omega Squared (ω²) | -0.259085 |

The ANOVA results for the statistical analysis, including the F-statistic, p-value, Omega Squared, and Least Significant Difference (LSD) from post hoc tests, are summarized in Table 21. These results provide insight into the variance between groups and the effect size of the observed differences.

***Between Groups Effects***

Between-group effects measure the variability due to differences between the group means and reflect the effect of the independent variable. Within-group effects reflect the variability within each group. The sum of squares (SS) quantifies the total variation for both between-group and within-group effects. The mean square (MS) was the sum of squares divided by respective degrees of freedom and represents the average variation for each effect. Higher mean squares suggest greater variability attributed to the effect being measured. (Creswell et al., 2023, p. 268).

**Table 22**

*ANOVA Between Groups Effects*

|  |  |
| --- | --- |
| Group A | Group B |
| Measure | Value |
| Sum of Squares | 0.058974 |
| Mean Square | 0.019658 |
| Degrees of Freedom | 3.000000 |

The table below presents the ANOVA results for the comparison between Group A and Group B. Table 22 shows the sum of squares, mean square, and degrees of freedom for the between-groups effects, highlighting the variance observed between these two groups.

***Within Group Effects***

Within-group effects represent the variability within each group and reflect differences within the same group. A higher mean square (MS) suggests more variation within the groups unrelated to the independent variable.

**Table 23**

*ANOVA Within Groups Effects*

|  |  |
| --- | --- |
| Measure | Value |
| Sum of Squares | 1.633333 |
| Mean Square | 0.181481 |
| Degrees of Freedom | 9.000000 |

A p-Value greater than the alpha level (0.05) indicates insufficient evidence to reject the null hypothesis and suggests no statistically significant difference between the group means.

A low (< 1) F-Statistic suggests little evidence that the group means are different. A large (> 0.14) Omega-Squared (ω²) statistic indicates a large effect size and suggests the independent variable has a strong effect on the dependent variable (Creswell et al., 2023, p. 268).

**Hypothesis 4**

**Ha4:** A statistically significant association exists between interview selection rates and genders among candidates in STEM fields who do not use Generative AI for resume creation or enhancement (RQ4).

**H04:** No statistically significant association exists between interview selection rates and genders among candidates in STEM fields who do not use Generative AI for resume creation or enhancement (RQ4).

The data set for hypothesis four was those who applied for a job in STEM in the past 12 months, who did not use Generative AI and who were selected for an interview. The categorical variables compared are the distribution of gender and interview selection. The sample size was n=13. The discussion below describes the statistical analysis for the ANOVA analysis for Hypothesis 4.

**Table 24**

*2x2 Chi-Squared Matrix for Hypothesis 4*

|  |  |
| --- | --- |
|  | Do not use Gen AI |
|  | **Selected** | **Not Selected** |
| Group 1 Male  | 8 | 2 |
| Group 2 Female | 3 | 0 |
| Total | 11 | 2 |

A Chi-Squared test of independence was conducted to examine the association between two categorical variables by comparing observed frequencies in a contingency table with expected frequencies derived under the hypothesis that the variables are independent. (MacFarland et al., 2016, p. 80; Creswell et al., 2023, p. 268).

**Table 25**

*Chi-Squared Analysis for Hypothesis 4*

|  |  |
| --- | --- |
| Statistic | Value |
| Chi-Squared Statistic | 0.7091 |
| Degrees of Freedom | 1.0000 |
| p-value | 0.39974577 |

The Chi-Squared procedure suggests no statistically significant association between Male, Selection and did not use Gen AI and Female, Selection and did not use Gen AI at an alpha level of 0.050 (p = 0.39974577).

Fisher's Exact Test suggests no statistically significant association between Male, Selection and did not use Gen AI and Female, Selection and did not use Gen AI at an alpha level of 0.050 (p = 0.39974577).

**Hypothesis 5**

The discussion below describes the Chi-Squared analysis for hypothesis five.

**H05:** No statistically significant association exists between interview selection rates and genders among candidates in STEM fields who use Generative AI for resume creation or enhancement (RQ5).

**Ha5:** A statistically significant association exists between interview selection rates and genders among candidates in STEM fields who use Generative AI for resume creation or enhancement (RQ5).

The data set was those who have looked for a job in STEM in the past 12 months, those who have used Generative AI and those who were selected for an interview. The categorical variables examined are gender distribution and interview selection. The sample size examined Is n=16.

**Table 26**

*2x2 Chi-square Matrix for Hypothesis 5*

|  |
| --- |
| Used Gen AI |
|   | **Selected** | **Not Selected** |
| Group 1 - Male | 6 | 2 |
| Group 2- Female | 6 | 2 |
| Total | 12 | 4 |

A Chi-Squared test of independence was conducted to examine the association between two categorical variables by comparing observed frequencies in a contingency table with expected frequencies derived under the hypothesis that the variables are independent. (MacFarland et al., 2016, p. 80; Creswell et al., 2023, p. 268).

**Table 27**

*Chi-Squared Analysis for Hypothesis 5*

|  |  |
| --- | --- |
| Statistic | Value |
| Chi-Squared Statistic | 0.0000 |
| Degrees of Freedom | 1.0000 |
| p-value | 1.00000000 |
| Fisher's Exact p-Value | 1.00000000 |

A Chi-squared statistic of 0. 0 and p-value of 1.0, which means the samples are identical. Therefore, the null hypothesis was rejected. Implying that use of Generative AI does not change the outcome for genders being selected for interviews.

**Hypothesis 6**

The following discussion examines hypothesis six.

**H06:** No statistically significant difference exists in the perceived accuracy of AI-generated resume content and racial groups among candidates in STEM fields? (RQ6).

**Ha6:** A statistically significant difference exists in the perceived accuracy of AI-generated resume content between racial groups among candidates in STEM fields? (RQ6).

The data set for hypothesis six was those who used Generative AI for the creation, revision or enhancement of their resumes and they categorized the results of Generative AI as accurate or not accurate. The categorical variables examined are racial distribution and binary perception of accuracy. The sample size was n=16.

An ANOVA (Analysis of Variance) procedure was applied to the sample to examine whether a statistically significant difference exists between the means of 5 unrelated groups (Creswell et al., 2023, p. 268).

**Table 28**

*Dependent Variable Summary by Independent Variable Groups*

|  |  |  |
| --- | --- | --- |
| Group | n | Mean(x̄) |
| White | 5 | 1.4000 |
| Black | 6 | 1.3333 |
| Asian | 3 | 1.0000 |
| 2 or More | 1 | 1.0000 |
| Other | 1 | 2.0000 |

***ANOVA Statistics***

**Table 29**

*ANOVA Results Statistics*

|  |  |
| --- | --- |
| Statistic | Value |
| F-Statistic | 0.981497 |
| p-Value | 0.45661248 |
| Omega Squared (ω²) | -0.004647 |
| LSD (post hoc) | 0.45661248 |

***Pairwise Comparisons for Significant Differences***

Pairwise comparisons in ANOVA were used to identify which specific groups differ significantly from each other after finding a significant overall effect by comparing the means of each pair of groups to determine where the differences lie (Creswell et al., 2023, p. 268).

**Table 30**

*Pairwise Comparisons for Significant Differences*

|  |  |
| --- | --- |
| Group A | Group B |
| Asian | Other (Native American) |
| Two or More | Other (Native American) |

In the pairwise comparison between Asian and Other (Native American), the significant difference was all of the Asians found Generative AI accurate, and none of the Native American respondents found Generative AI accurate. The sample size for Native Americans was one. The same finding was true for two or more races where the sample size was also one. There was a pairwise statistical difference in the perception of Generative AI’s accuracy, however, the difference was influenced by the small sample size.

***Between Groups Effects***

Between-group effects measure the variability due to differences between the group means and reflect the effect of the independent variable. Within-group effects reflect the variability within each group. The sum of squares (SS) quantifies the total variation for both between-group and within-group effects. The mean square (MS) was the sum of squares divided by respective degrees of freedom and represents the average variation for each effect. Higher mean squares suggest greater variability attributed to the effect being measured (Creswell et al., 2023, p. 268).

**Table 31**

*ANOVA Between Groups Effects*

|  |  |
| --- | --- |
| Measure | Value |
| Sum of Squares | 0.904167 |
| Mean Square | 0.226042 |
| Degrees of Freedom | 4.000000 |

***Within Group Effects***

Within-group effects represent the variability within each group and reflect differences within the same group. A higher mean square (MS) suggests more variation within the groups unrelated to the independent variable.

**Table 32**

*ANOVA Within Groups Effects*

|  |  |
| --- | --- |
| Measure | Value |
| Sum of Squares | 2.533333 |
| Mean Square | 0.230303 |
| Degrees of Freedom | 11.000000 |

**Hypothesis 7**

**H07:** No statistically significant difference exists in the perceived accuracy of AI-generated resume content between genders among candidates in STEM fields (RQ7).

**Ha7:** A statistically significant difference exists in the perceived accuracy of AI-generated resume content between genders among candidates in STEM fields (RQ7).

The data set for hypothesis seven was those who used Generative AI for the creation, revision or enhancement of their resumes and their perception of the accuracy of the Generative AI model in describing their qualifications. The categorical variables examined are the gender distribution (independent) and the perception of accuracy (dependent). The sample size was n=20.

The Mann-Whitney U test was conducted to determine if there were statistically significant differences between the means of men and women. (Creswell et al., 2023, p. 255).

**Table 33**

*Mann Whitney U Statistics*

|  |  |
| --- | --- |
| Measure | Value |
| Group 1 (Male) Mean | 1.3750 |
| Group 2 (Female) Mean | 1.2500 |
| Mann Whitney U | 59.0000 |
| Z-Score | 2.8356 |
| p-value | 0.00457444 |

***Post-Hoc Procedures***

The Mann-Whitney U post-hoc statistics are found in Table 34. The Mann-Whitney U test is a non-parametric statistical test used to compare differences between two independent groups with data which is not normally distributed. In post hoc analysis, additional statistics are calculated to further interpret the results. For this analysis Boneferroni Correction Alpha is set at 0.050, which adjusts the significance level to reduce the likelihood of Type 1 errors due to multiple comparisons. Cohen’s d (effect size) is 0.2546, indicating a small to medium effect size. Cohen’s d quantifies the difference between two group means in terms of standard deviation and helps in understanding the magnitude of the effect. The power of the test is 0.9956, suggesting a high probability (99.56%) of correctly rejecting the null hypothesis when it is false. This indicates that the test is highly reliable in detecting true effects in the data. These statistics provide a comprehensive understanding of the differences between the groups and the reliability of the findings (Creswell et al., 2023, p. 255).

**Table 34**

*Mann-Whitney U Post Hoc Statistics*

|  |  |
| --- | --- |
| Statistic | Value |
| Bonferroni Correction Alpha | 0.050000 |
| Cohen's d (effect size) | 0.254588 |
| Power | 0.9956 |

A two-tailed Mann-Whitney U procedure was applied to the samples assuming equal variances to test the null hypothesis that the difference in means the dependent variable between Male and Female were not equal to zero. The means for the groups Male and Female were 1.3750 and 1.2500 respectively. The p-value of 0.00457444 suggests a statistically significant difference between the groups' distributions at a 0.05 alpha level. The null hypothesis was rejected. (Creswell et al., 2023, p. 255).

## Summary of Findings

This section is a summary of the findings related to each research question, which was answered based on each associated hypothesis test. Whether each null hypothesis was rejected, the strength of each statistical measure, and relevant post-hoc statistics are considered.

**Research Question 1**

* **RQ1:** What associations exist in interview selection rates between those who use Generative AI for resume creation or enhancement and those who do not among candidates in STEM fields?

**Table 35**

*Research Question 1 Statistics Summary*

|  |  |
| --- | --- |
| RQ 1 Summary Table | Value |
| Chi-Squared Statistic | 0.4041 |
| Degrees of Freedom | 1.0 |
| Significance Level  | 0.05 |
| Critical Value approx | 3.841 |
| d p-value | 0.52496666 |

***Research Question 1 Chi-Squared Findings***

The Chi-Squared procedure suggests no statistically significant association between Used AI and Did not use Gen AI at an alpha level of 0.050 (p = 0.52496666).

Fisher's Exact Test suggests no statistically significant association between Used AI and Did not use Gen AI at an alpha level of 0.050 (p = 0.52496666).

Given the Chi-Squared statistic 0.4041, with degrees of freedom of 1and significance level of 0.05 the critical value from the Chi-Squared distribution table was approximately 3.841. The Chi-squared statistic was much lower than the critical value. This suggests that the observed frequencies do not significantly differ from the expected frequencies and any difference could be due to random chance rather than a true effect. No meaningful associations existed in interview selection rates between those who use Generative AI for resume creation or enhancement and those who do not among candidates in STEM fields. (MacFarland et al., 2016, p. 80; Creswell et al., 2023, p. 268)

**Research Question 2**

* RQ2: What differences exist in interview selection rates between racial groups among candidates in STEM fields who use Generative AI for resume creation or enhancement?

A summary of the ANOVA statistics is depicted in Table 36.

**Table 36**

*Research Question 2 Statistics Summary*

|  |  |
| --- | --- |
| Statistic | Value |
| F-Statistic | 1.307377 |
| p-Value | 0.32644373 |
| Omega Squared (ω²) | 0.071361 |
| LSD (post hoc) | 0.32644373 |

***Research Question 2 ANOVA Findings***

There was a significant difference between the White group and the Asian, Two or more and other (South Asian) groups. There were a significantly higher percentage of whites who use Generative AI and were selected for interviews and Asian (0%), Two or More (9.1%) and Other (9.1%) who were selected for interviews. While the numbers were small for the Asian, Two or more races, a much higher percentage of whites were selected or interviews with the use of Generative AI.

A p-Value greater than the alpha level (0.05) indicates insufficient evidence to reject the null hypothesis and suggests no statistically significant difference between the group means. An ANOVA F-statistic of 1.30 suggests variability between the group means were slightly greater than the variance within the groups. Thus, the differences among group means were not statistically significant. In conclusion, a p-Value greater than the alpha level (0.05) indicates insufficient evidence to reject the null hypothesis and suggests no statistically significant difference between the group means of racial groups who use Generative AI and being selected for an interview. No meaningful differences exist in interview selection rates between racial groups among candidates in STEM fields who use Generative AI for resume creation or enhancement.

**Research Question 3**

* RQ3: What differences exist in interview selection rates between racial groups among candidates in STEM fields who do not use Generative AI for resume creation or enhancement?

**Table 37**

*Research Question 3 Statistics Summary*

|  |  |
| --- | --- |
| p-value | 0.953094 |
| ANOVA F-Statistic | 0.108320 |
| Omega Squared (ω²) | -0.259085 |
| p-value | 0.953094 |

***Research Question 3 ANOVA Findings***

An ANOVA F-statistic of 0.108320 suggests variability within each group is much greater than the variability between the group means. Thus, the differences between group means are not statistically significant. In conclusion, a p-Value greater than the alpha level (0.05) indicates insufficient evidence to reject the null hypothesis and suggests no statistically significant difference between the group means of racial groups who did not use Generative AI and being selected for an interview. No meaningful differences exist in interview selection rates between racial groups among candidates in STEM fields who do not use Generative AI for resume creation or enhancement.

**Research Question 4**

* RQ4: What association exists in interview selection rates between genders among candidates in STEM fields who do not use Generative AI for resume creation or enhancement?

**Table 38**

*Research Question 4 Statistics Summary*

|  |  |
| --- | --- |
| Statistic | Value |
| Chi-Squared Statistic | 0.7091 |
| Degrees of Freedom | 1.000 |
| Significance Level  | 0.05 |
| Critical Value | 3.814 |
| Chi-Squared p-value | 0.39974577 |

***Research Question 4 Chi-Squared Findings***

The Chi-Squared procedure suggests no statistically significant association between Gender groups that did not use Generative AI, and I was selected for an interview for the position for which I applied. at an alpha level of 0.050 (p = 0.39974577).

Fisher's Exact Test suggests no statistically significant association between Gender groups that did not use Generative AI, and I was selected for an interview for the position for which I applied. at an alpha level of 0.050 (p = 0.39974577).

The Chi-Squared procedure suggests no statistically significant association between Gender Groups did not use Generative AI and I was selected for an interview at an alpha level of 0.050 (0.39974577). Given the Chi-Squared statistic of 0.7091, with degrees of freedom 1 and significance level of 0.05 the critical value from the Chi-Squared distribution table is approximately 3.814. The Chi-Squared statistic is much lower than the critical value. This suggests that the observed frequencies do not significantly differ from the expected frequencies and any difference could be due to random chance rather than a true effect. No meaningful association existed in interview selection rates between genders among candidates in STEM fields who do not use Generative AI for resume creation or enhancement. (MacFarland et al., 2016, p. 80; Creswell et al., 2023, p. 268).

**Research Question 5**

* RQ5: What association exists in interview selection rates between genders among candidates in STEM fields who use Generative AI for resume creation or enhancement?

**Table 39**

*Research Question 5 Statistics Summary*

|  |  |
| --- | --- |
| Statistic | Value |
| Chi-Squared Statistic | 0.0 |
| Degrees of Freedom | 1.000 |
| Significance Level  | 0.05 |
| Chi-Squared p-value | 1.0 |

***Research Question 5 Chi-Squared Findings***

The Chi-Squared procedure suggests no statistically significant association between Male, Selection and did not use Gen AI and Female, Selection and did not use Gen AI at an alpha level of 0.050 (p = 1.0).

Fisher's Exact Test suggests no statistically significant association between Male, Selection and did not use Gen AI and Female, Selection and did not use Gen AI at an alpha level of 0.050 (p = 1.0). No meaningful association existed in interview selection rates between genders among candidates in STEM fields who use Generative AI for resume creation or enhancement. (MacFarland et al., 2016, p. 80; Creswell et al., 2023, p. 268).

**Research Question 6**

* RQ6: What differences exist in the perceived accuracy of AI-generated resume content between racial groups among candidates in STEM fields?

**Table 40**

*Research Question 6 Statistics Summary*

|  |  |
| --- | --- |
| Statistic | Value |
| F-Statistic | 0.981497 |
| p-Value | 0.45661248 |
| Omega Squared (ω²) | -0.004647 |

***Research Question 6 ANOVA Findings***

A p-Value greater than the alpha level (0.05) indicates insufficient evidence to reject the null hypothesis and suggests no statistically significant difference between the group means. A low (< 1) F-Statistic suggests little evidence that the group means are different. A large (> 0.14) Omega-Squared (ω²) statistic indicates a large effect size and suggests the independent variable has a strong effect on the dependent variable. No meaningful differences existed in the perceived accuracy of AI-generated resume content between racial groups among candidates in STEM fields.

**Research Question 7**

* RQ7: RQ7: What differences exist in the perceived accuracy of AI-generated resume content between genders among candidates in STEM fields?

***Research Question 7 Mann-Whitney U Findings***

**Table 41**

*Research Question 7 Statistics Summary*

|  |  |
| --- | --- |
| Statistic | Value |
| Group 1 (Male) Mean | 1.3750 |
| Group 2 (Female) Mean | 1.2500 |
| Mann Whitney U | 59.0000 |
| Z-Score | 2.8356 |
| p-value | 0.00457444 |

A two-tailed Mann-Whitney U procedure was applied to the samples assuming equal variances to test the null hypothesis that the difference in means the dependent variable between Male and Female were not equal to zero. The means for the groups Male and Female were 1.3750 and 1.2500 respectively.

The p-value of 0.00457444 suggests a statistically significant difference between the groups' distributions at a 0.05 alpha level. The null hypothesis was rejected. The female respondents viewed Generative AI as significantly more accurate than the male respondents. 62% of female respondents viewed Generative AI as accurate, while 62% of male respondents viewed Generative AI as inaccurate. The suggests an association difference in accuracy of Generative AI’s use between male and females.

## Conclusion

Data were collected and analyzed from 35 respondents using Survey Monkey’s on-line platform over 5 weeks, June - July 2024. 83% of the respondents were college educated and applied for jobs in STEM fields. Due to the Snowball sampling technique, the respondents primarily represented careers in the computer science field which would be early adopters in the use of Generative AI. 76% of the respondents were from minorities and 43% were women. The purpose of this study is to examine the association in interview selection rates among marginalized groups in STEM fields according to the use of Generative AI for resume creation or enhancement. Consistently, minorities used Generative AI in larger percentages than the majority sample, comprising white and male applicants. Of the minority sample that used AI, 42% of blacks and 25% of Asians were selected for interviews. Only 17% of those selected for interviews who used Generative AI were white because fewer reported using the technology. 45% of whites did not use Generative AI compared to only 36% of blacks. The same trend followed for men and women. More women used Generative AI, but the interview selection was equal. The accuracy was perceived to be greater by the minority sample and women than by the non-diverse sample and men.

The research questions were analyzed using Chi-Squared, ANOVA or Mann-Whitney U. The statistical analysis procedure for Chi-squared, and ANOVA for each of the research questions suggests no statistically significant association between gender and race who used Generative AI and were selected for interviews. There was a significant association between using Generative AI and accuracy between gender groups using the Mann-Whitney U statistical analysis procedure. 62% of Female users of Generative AI found it accurate, while 62% of males found it inaccurate.

In a study on the impact of Generative AI on historically marginalized populations in STEM this study suggests that Generative AI can provide greater value for females developing resume material than men. While there was a significant difference found in racial accuracy perception between Asians, who found Generative AI 100% accurate, and two or more races and Native Americans who found it inaccurate, the sample size was too small to draw conclusions. In all cases, the ANOVA F-statistic suggests variability within each group is greater than the variability between the group means. Thus, the differences between racial and gender group means are not statistically significant. The Mann-Whitney U test suggests statistically significant evidence to reject the null hypothesis in hypothesis 7. There is evidence to suggest a difference between the perception of Generative AI’s accuracy between men and women who use Generative AI to create or enhance their cover letter or resume. Women perceive Generative AI to be more accurate than men.

In conclusion, a p-Value greater than the alpha level (0.05) indicates insufficient evidence to reject the null hypothesis for six of seven of the research questions. The perception of accuracy when using Generative AI was statistically significant between men and women, and the null hypothesis was rejected. While the number of respondents was 35, subdividing into smaller samples reduced the statistical significance. The survey will remain open and additional data collection will continue given the relatively recent release of Generative AI and the sample's increasing use preparing pre-employment documentation using Generative AI.

## References

Creswell, J. W., & Creswell, J. D. (2023). Research design qualitative, quantitative, and mixed methods approaches. Sage.

MacFarland, T.W. Yates, J.M. (2016) Introduction to nonparametric statistics for the biological sciences using R, Springer International Publishing. DOI 10.1007/978-3-319-30634-6\_3.

Reichard, J. (2024). Practical Statistics for Social Research (PSSR). Omega Graduate School. <https://stats.ogs.edu>.