



HR Statistics 101

Part 2 of 2

Presented By:
Valentin Estevez, Ph.D.,
Dan (dk) Kuang, Ph.D.

PRESENTERS



Valentin Estevez, Ph.D.

Charles River Associates
vestevez@crai.com



Dan (dk) Kuang, Ph.D.

Resolution Economics / Berkshire Associates
dkuang@resecon.com

DISCLAIMER

- The information in this PowerPoint and the related webinar presentation is provided for general information purposes only and does **not constitute legal advice**. You should consult your legal counsel for advice regarding the applicability of the information presented to your specific situation.

RECAP STATS 101 PART 1

- HRIS Data
- Data Analysis Basics
- Data Visualization
- Data Comparisons

AGENDA—PART 2

Inferential Statistics

- Selection Rate Analysis
- Test of Proportions: Binomial
- Multiple Regression: Differences in Means
- Post-Analysis Diagnostic Reviews
 - Uncover reasons for unexplained differences

Recap on “Statistical Significance”

Outcome differences equivalent to 2+ standard deviations (“SDs”) are statistically significant in employment matters.

Statistical significance does not mean a difference exists.

- It indicates that a correlation between outcomes and group membership exists that requires further review.

Statistical significance does not prove discrimination.

- However, statistically significant differences can be used as evidence of discrimination in legal proceedings.

Statistical significance is different from practical significance.

- Unfortunately, no agreed-upon measure of practical significance exist. Context matters.

Recap on “Statistical Significance”

Why the need to examine the statistical results?

- False positives or Type I errors: **Detecting** a difference when there is not one
- False negatives or Type II errors: **Failing to detect** a difference when there is one

Robust data, including documentation, and careful review of the statistical results are best practices to minimize the incidence of Type I and Type II errors.

Outcome differences of any size are more likely to be over 2 SDs in larger databases.

Where do we start?

Fortunately, several well-understood and widely-used statistical techniques exist to analyze employment outcomes.

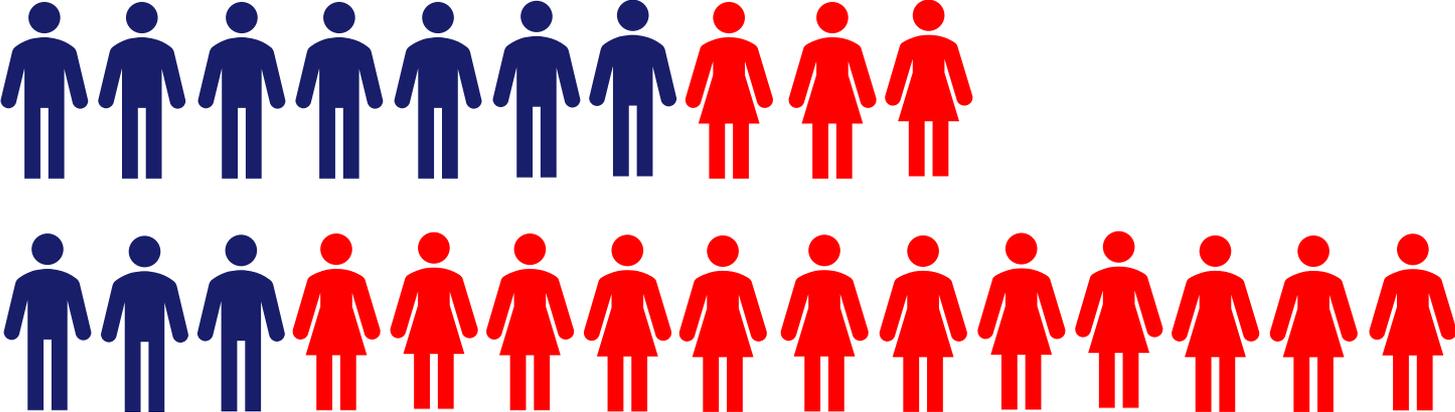
Some are more complex than others. But deep down, they operate similarly.

| Intro | Advanced |
|---|-----------------------------------|
| T-tests for average compensation comparisons | Multiple regression |
| Chi-square/Fisher's exact for selection rates | Logit/probit regression |
| Cohort/group analyses | Regression residual reviews |
| Scatter plots/data visualization | Scatter plots/data visualization* |
| * No matter the approach, <i>always look at your data.</i> | |

Selection Rates

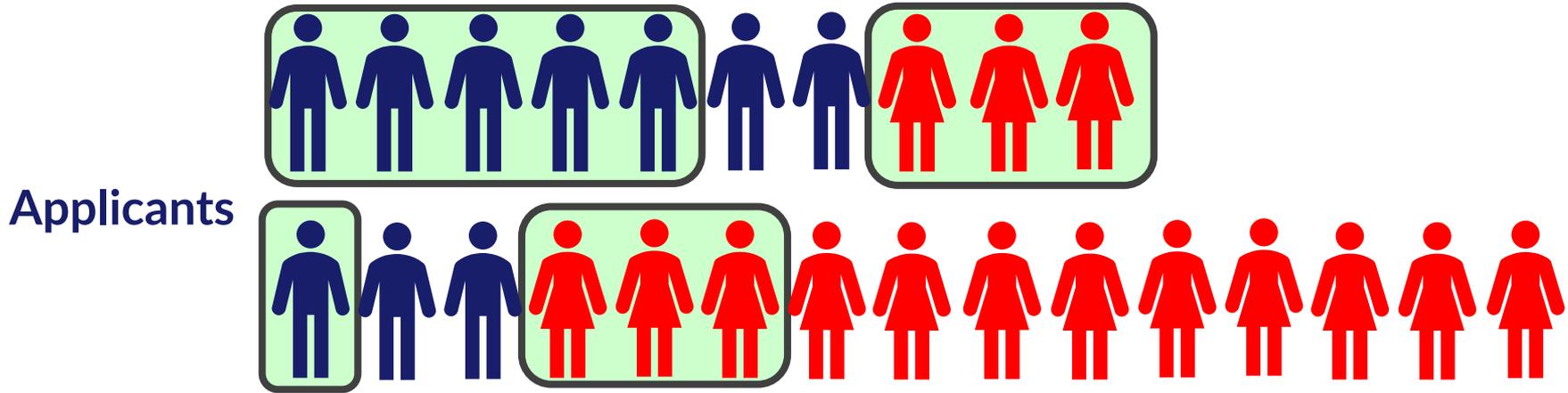
Thinking About Selections

Applicants



Each stick figure = 15 applications, or 375 total applications (225 female, 150 male)

Thinking About Selections



48% Selected (highlighted stick figures) = 180 selections (90 female, 90 male)

Female selection rate: 40% (= 90 / 225)

Male selection rate: 60% (= 90 / 150)

Female - Male selection rate difference: -20%

Thinking About Selections

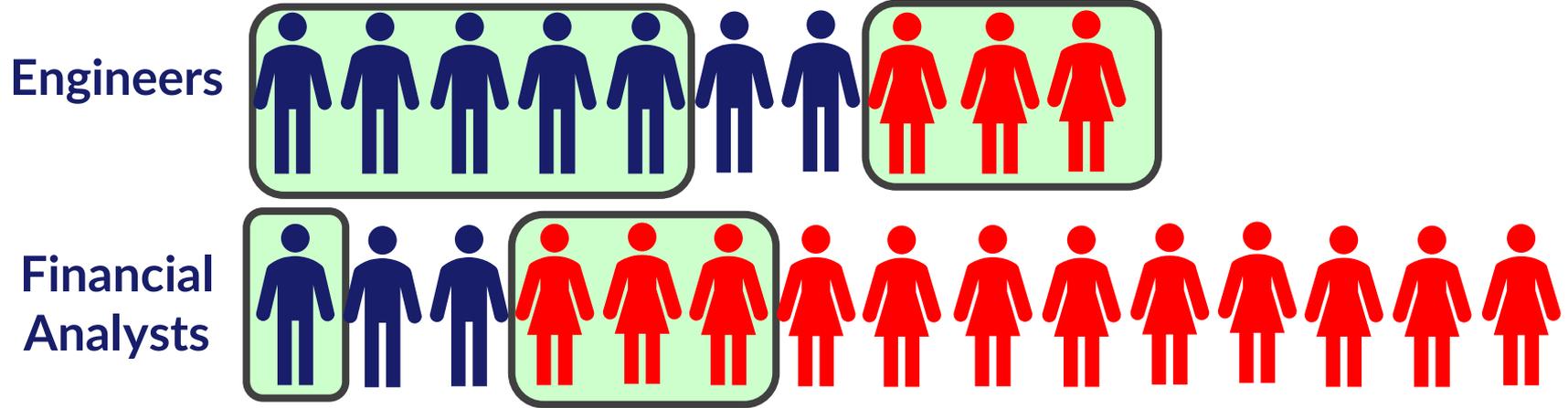
| Applications | Female Applicants | Actual Females Selections | Expected Female Selections | Actual Minus Expected Selections | SDs | Stat. Sig? |
|--------------|-------------------|---------------------------|----------------------------|----------------------------------|------|------------|
| All | 225 | 90 | 108 | -18.0 | -3.8 | Yes |

Expected Female Selections = Overall Selection Rate * Female Population

Expected Female Selections = 48% * 225 = 108

Difference in SD Units = -3.8; statistically significant (i.e., greater than 2 SDs)

Digging Deeper, Thanks to Data and Documents



Engineer applicant selection rate: 100% of Females, 71.4% of Males

Financial analyst applicant selection rate: 25% of Females, 33.3% of Males

A Different Picture Emerges

| Type | Female Applicants | Actual Females Selected | Expected Female Selected | Actual Minus Expected Selections | SDs | Stat. Sig? |
|-----------------------------------|-------------------|-------------------------|--------------------------|----------------------------------|------|------------|
| Engineers | 45 | 45 | 36.0 | 9.0 | 4.0 | Yes |
| Financial Analysts | 180 | 45 | 48.0 | -3.0 | -1.1 | No |
| All, <u>But Adjusting for Job</u> | 225 | 90 | 84.0 | 6.0 | 1.7 | No |

Note on sample sizes: All three comparisons would be over 2 SDs if there were 900 applications (4 times more data), while maintaining the same demographic composition, applications by job, and selection rates of the original example.

Test of Proportions

SELECTIONS RATE ... TEST OF PROPORTIONS

Comparing selection rates (e.g., testing, hiring, promotions, terminations) or HR decisions with a limited number of outcomes (performance reviews).

How to evaluate? 4/5ths or 80% rule of thumb

| | Pass | Fail |
|--------|------------|------------|
| Male | M_{pass} | M_{fail} |
| Female | F_{pass} | F_{fail} |

$$\text{Selection Ratio (Impact Ratio)} = \frac{\frac{F_{pass}}{F_{pass} + F_{fail}}}{\frac{M_{pass}}{M_{pass} + M_{fail}}}$$

TEST OF PROPORTIONS—WHEN TO USE...

- ...data is messy.
 - Selection Rate Analysis may not be reliable
- ...results need to be easily understood (e.g., executives, stakeholders)
 - Proportion comparisons are very INTUITIVE
- ...need a quick & dirty answer/insight
 - Reasonably robust method (to get a thumbnail answer)

TEST OF PROPORTIONS—BASICS

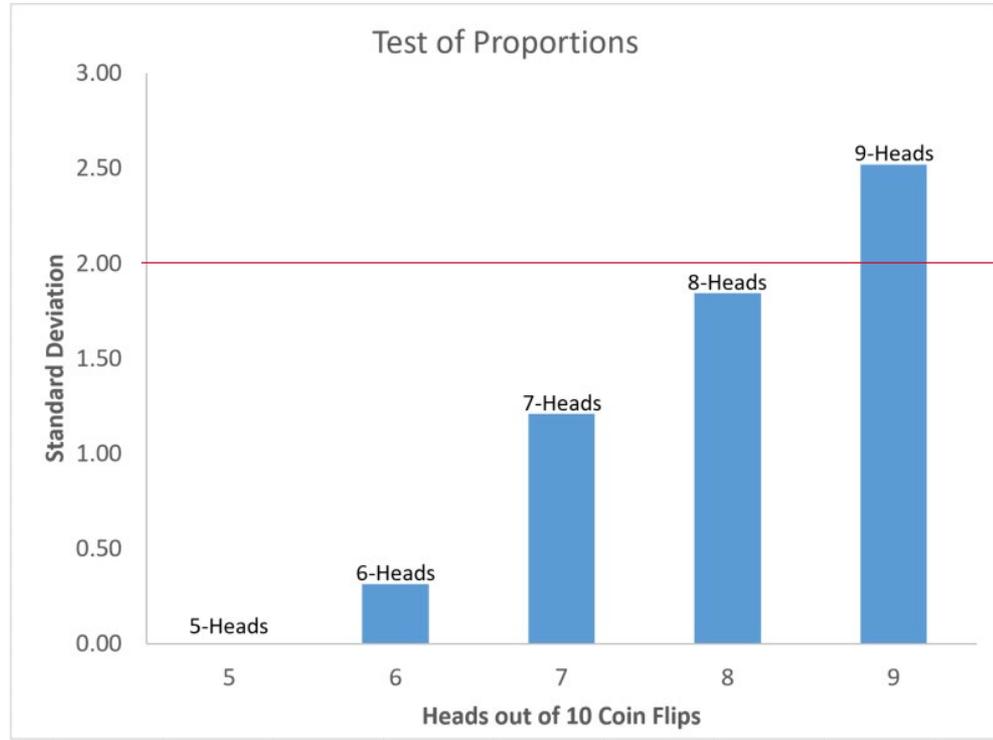
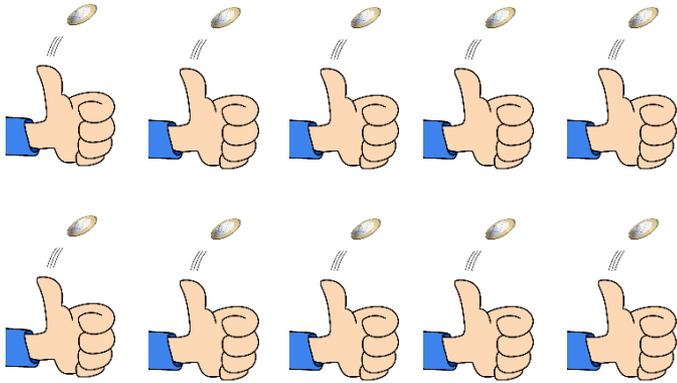
- **Q:** Are there differences in Proportions between Observed and Expected?



| Outcome | Expected (%) | Observed (%) |
|---------|--------------|--------------|
| Heads | 50% | 5 of 10 |
| Tails | 50% | 5 of 10 |

TEST OF PROPORTIONS—BASICS

- **Q:** Are there differences in Proportions between Observed and Expected?



TEST OF PROPORTIONS—SUPREME COURT

Hazelwood School District v. United States, 433 US 299 (1977)

- Workforce Composition Analysis
- Black Teachers in Hazelwood School District

| Year | Total | Obs (#) | Obs (%) |
|-----------|-------|---------|---------|
| 1972-1973 | 282 | 10 | 3.5% |
| 1973-1974 | 123 | 5 | 4.1% |

St. Louis County = 5.7%

| Year | Exp (#) | SD |
|-----------|---------|--------|
| 1972-1973 | 16 | < 2 SD |
| 1973-1974 | 7 | < 1 SD |

City of St. Louis = 15.4%

| Year | Exp (#) | SD |
|-----------|---------|-------|
| 1972-1973 | 43 | 5+ SD |
| 1973-1974 | 19 | 3+ SD |

TEST OF PROPORTIONS—MESSY APPLICANT DATA

Q: How do you analyze messy or incomplete Applicant Data for potential adverse impact in hires?

- Test of proportions can help evaluate for potential imbalance between applicants and hires.
- For example:

Applicants

| Groups | Counts(#) | Pct (%) |
|--------------|-----------|---------|
| Pink | 750 | 75% |
| Purple | 250 | 25% |
| Total | 1,000 | |

Hired

| Groups | Counts(#) | Pct (%) | <i>SD</i> |
|--------------|-----------|---------|-----------|
| Pink | 10 | 50% | 2.39 |
| Total | 20 | | |

Multiple Regression

Mean or Average Comparisons – Starting Point

Comparing mean outcomes for base pay, incentive compensation, numerical scores, etc.

How to evaluate? Relative to average raises, external benchmarks. Are there legitimate explanations for the observed differences?

| Analysis Group | Female Average Annual Salary | Male Average Annual Salary | Gender Difference | Gender Pay Ratio |
|-----------------------|---|---------------------------------------|--------------------------|-------------------------|
| Analysis Group 01 | \$ 71,112 | \$ 66,459 | -\$ 5,347 | 107.0% |
| Analysis Group 02 | \$ 96,588 | \$ 96,906 | -\$ 318 | 99.7% |
| ... | | | | |
| Analysis Group 10 | \$ 151,462 | \$ 149,155 | \$ 2,307 | 101.5% |
| Analysis Group 11 | \$ 161,605 | \$ 167,405 | -\$ 5,800 | 96.5% |
| Analysis Group 12 | \$ 206,665 | \$ 212,843 | -\$ 6,178 | 97.1% |

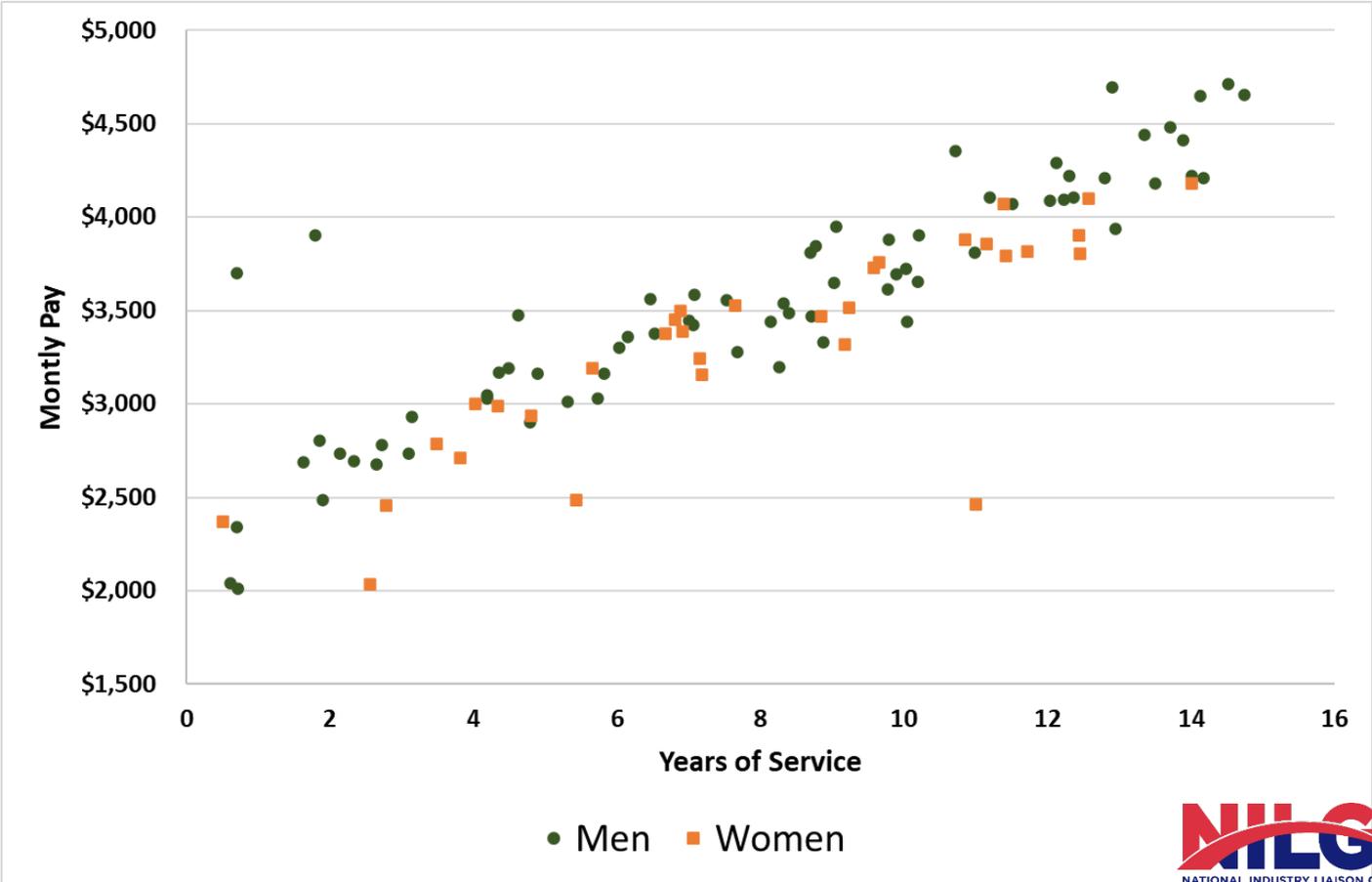
Multiple Regression: Estimating the Impact of Pay Factors

Multiple regression estimates the **average** relationship between the dependent variable (e.g., pay) and the independent variables in the model (e.g., years of service).

One of the most commonly used statistical techniques.

- Well-understood and accepted.
- Standard in statistical applications. Add-ins are also available (e.g., Analysis ToolPak in Excel).
- Flexible.
- Incorporates powerful diagnostic tools for deeper dives.

Multiple Regression: Estimating the Impact of Pay Factors



Multiple Regression: Estimating the Impact of Pay Factors



Multiple Regression: Where Do Differences Come From?



Prediction Error: Variation Unaccounted For By Pay Factors



Are There Differences Correlated with Group Membership?

| Analysis Group | Emps | Male | White | Gender Comparison | | | Race Comparisons Relative to Whites | | | | | | | | |
|-------------------|-------|-------|-------|-------------------|------------|------------|-------------------------------------|------------|------------|--------|------------|------------|-------|------------|------------|
| | | | | Fem. | Pct. Diff. | Std. Devs. | Afr. Amer. | Pct. Diff. | Std. Devs. | Asians | Pct. Diff. | Std. Devs. | Hisp. | Pct. Diff. | Std. Devs. |
| Job Family | | | | | | | | | | | | | | | |
| Job Family 01 | 3,068 | 2,467 | 1,242 | 601 | - 4.2% | - 3.27 | 49 | - 1.9% | - 1.33 | 1,718 | - 0.8% | - 0.18 | 59 | 0.8% | 0.06 |
| Job Family 02 | 300 | 91 | 212 | 209 | - 1.8% | - 0.86 | 4 | 0.7% | 0.13 | 59 | 3.5% | 1.43 | 25 | 2.8% | 0.81 |
| Job Family 03 | 368 | 317 | 298 | 51 | 2.1% | 0.69 | 10 | 3.4% | 2.13 | 29 | 3.9% | 1.21 | 31 | - 0.7% | - 0.23 |
| Job Family 04 | 82 | 42 | 57 | 40 | - 1.1% | - 0.32 | 3 | - 0.2% | - 0.30 | 10 | 1.4% | 0.30 | 12 | - 0.4% | - 0.78 |

| Job Title | Emps | Male | White | Fem. | Pct. Diff. | Std. Devs. | Afr. Amer. | Pct. Diff. | Std. Devs. | Asians | Pct. Diff. | Std. Devs. | Hisp. | Pct. Diff. | Std. Devs. |
|---------------------|-------|------|-------|------|------------|------------|------------|------------|------------|--------|------------|------------|-------|------------|------------|
| Software Engineer | 430 | 345 | 174 | 84 | 2.0% | 2.30 | 7 | - 0.3% | - 0.19 | 241 | - 0.1% | - 0.03 | 8 | 0.1% | 0.01 |
| Aerospace Engineer | 1,043 | 839 | 422 | 204 | - 8.0% | - 3.04 | 17 | - 0.7% | - 0.45 | 584 | - 0.3% | - 0.06 | 20 | 0.3% | 0.02 |
| Mechanical Engineer | 250 | 150 | 189 | 100 | 2.0% | 1.44 | 6 | 1.3% | 0.65 | 33 | 2.9% | 0.98 | 23 | 0.6% | - 0.07 |
| | | | | | | | | | | | | | | | |
| Technicians | 381 | 285 | 176 | 96 | - 1.4% | - 0.62 | 8 | - 0.4% | - 0.40 | 181 | 1.2% | 0.26 | 17 | - 0.3% | - 0.70 |

| Level | Emps | Male | White | Fem. | Pct. Diff. | Std. Devs. | Afr. Amer. | Pct. Diff. | Std. Devs. | Asians | Pct. Diff. | Std. Devs. | Hisp. | Pct. Diff. | Std. Devs. |
|----------|-------|-------|-------|------|------------|------------|------------|------------|------------|--------|------------|------------|-------|------------|------------|
| Level 01 | 1,319 | 1,061 | 534 | 258 | - 6.0% | - 2.21 | 21 | - 0.8% | - 0.57 | 739 | - 0.3% | - 0.08 | 25 | 0.3% | 0.03 |
| Level 02 | 99 | 86 | 80 | 14 | - 5.0% | - 1.30 | 3 | 0.9% | 0.58 | 8 | 1.0% | 0.33 | 8 | - 0.2% | - 0.06 |
| Level 03 | 250 | 150 | 189 | 100 | 2.0% | 1.44 | 6 | 1.3% | 0.65 | 33 | 2.9% | 0.98 | 23 | 0.6% | - 0.07 |

Statistically Significant and Negative.

Statistically Significant and Positive.

Diagnostic Reviews

Multiple Regression: Several Diagnostic Tools Available

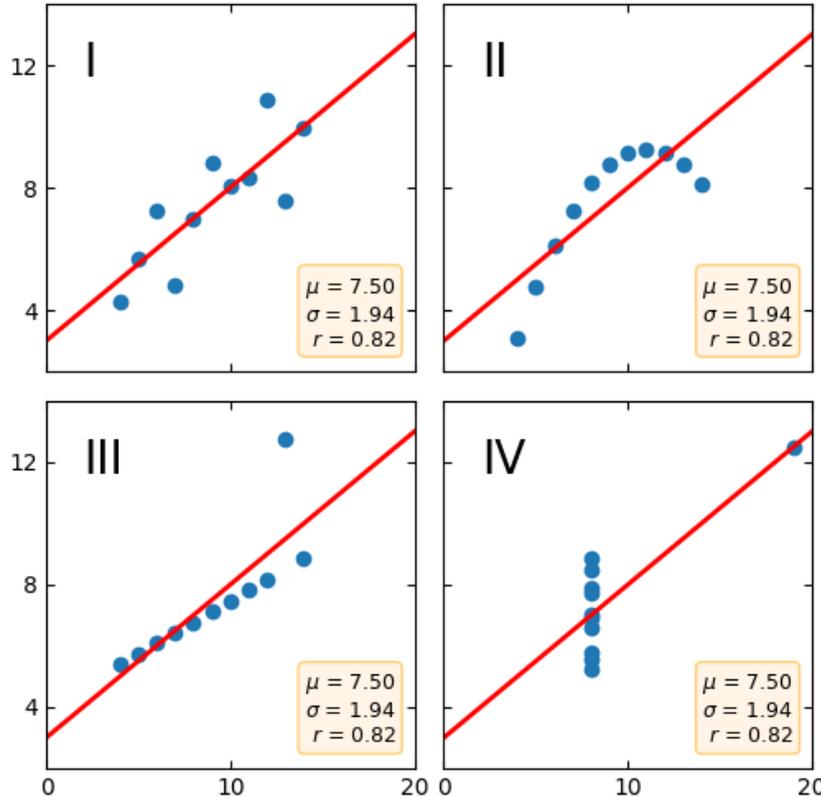
- **Outlier individuals**
 - Individuals whose compensation differs from what's expected, *given the model and available data.*
- **Outlier groups of individuals**
 - Subsets of the data accounting with above-average unexplained differences, *given the model and available data..*
- **Influential (high-leverage) statistics**
 - Individuals whose data push or pull the regression line.
 - May not be an outlier.

Multiple Regression: Several Diagnostic Tools Available

| Job | % Diff From This Job | ID | Sex | Race | Annual Salary | Predicted Annual Salary | Pct. Pred. Error | Location | Time in Job | Time in Other Jobs | Performance Ratings |
|--------------------|----------------------|------|-----------|-----------|---------------|-------------------------|------------------|----------|-------------|--------------------|---------------------|
| Software Engineer | -26.7% | 3744 | F | B | \$ 59,546 | \$ 80,014 | -25.3% | Loc 65 | 2.8 | 0.0 | Score 8 |
| | | 1780 | M | B | \$ 64,508 | \$ 81,170 | -20.2% | Loc 65 | 0.0 | 2.9 | Score 8 |
| | | 2135 | F | W | \$ 71,992 | \$ 81,935 | -11.8% | Loc 65 | 0.4 | 2.6 | Score 2 |
| | | 4245 | F | A | \$ 69,378 | \$ 78,394 | -11.2% | Loc 65 | 0.8 | 1.4 | Score 8 |
| | | 4262 | M | A | \$ 84,765 | \$ 77,324 | 10.0% | Loc 65 | 0.0 | 3.2 | Score 3 |
| | | 3377 | F | A | \$ 90,706 | \$ 82,574 | 10.3% | Loc 65 | 0.9 | 0.9 | Score 3 |
| | | 1128 | M | A | \$ 86,555 | \$ 78,343 | 10.9% | Loc 65 | 9.0 | 1.9 | Score 3 |
| | | 2219 | M | A | \$ 91,023 | \$ 82,368 | 10.9% | Loc 65 | 0.1 | 0.8 | Score 8 |
| | | 171 | M | W | \$ 86,183 | \$ 76,963 | 12.4% | Loc 65 | 0.2 | 2.2 | Score 3 |
| 559 | M | B | \$ 97,176 | \$ 83,294 | 17.1% | Loc 65 | 6.2 | 2.5 | Score 8 | | |
| Aerospace Engineer | -18.1% | 1323 | F | A | \$ 102,936 | \$ 129,743 | -20.4% | Loc 65 | 1.7 | 3.8 | Score 7 |
| | | 4001 | F | A | \$ 106,283 | \$ 127,039 | -16.0% | Loc 65 | 0.3 | 2.8 | Score 3 |
| | | 2873 | M | H | \$ 139,558 | \$ 136,536 | 2.6% | Loc 65 | 0.5 | 1.4 | Score 3 |
| | | 3797 | M | A | \$ 132,118 | \$ 129,033 | 2.8% | Loc 65 | 0.0 | 4.1 | Score 7 |
| Civil Engineer | -12.6% | 2976 | F | A | \$ 95,780 | \$ 115,631 | -16.8% | Loc 60 | 0.2 | 1.4 | Score 3 |
| | | 153 | F | A | \$ 102,785 | \$ 113,245 | -8.9% | Loc 60 | 0.7 | 0.0 | Score 3 |
| | | 97 | M | A | \$ 129,398 | \$ 122,351 | 6.2% | Loc 13 | 8.7 | 1.6 | Score 3 |
| | | 1874 | M | W | \$ 117,003 | \$ 109,542 | 7.2% | Loc 09 | 2.2 | 0.8 | Score 3 |

Multiple Regression: Always Examine Your Data!

All Four Sets of Data Points Yield the Same Statistics!

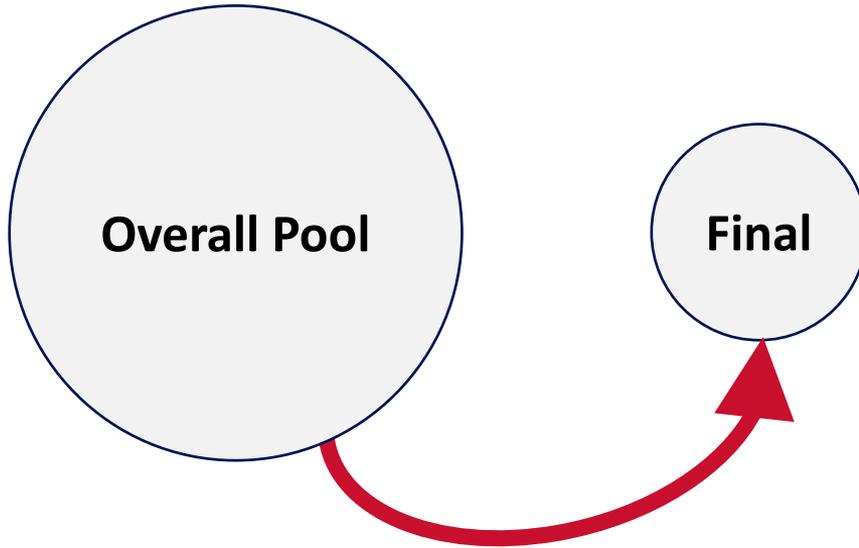


← Influential (High-Leverage) Data Point That is Not an Outlier

Source: https://matplotlib.org/stable/gallery/specialty_plots/anscombe.html

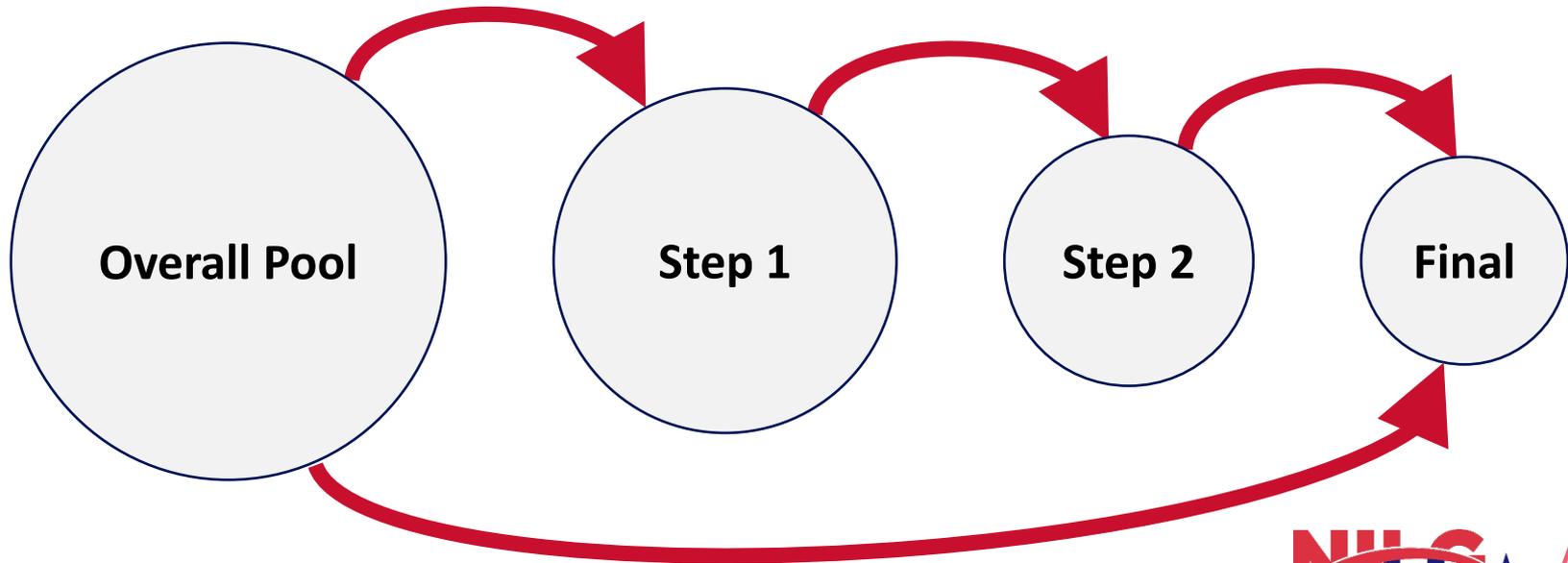
SELECTION RATE: DIAGNOSTICS

- Selection Steps Analysis



SELECTION RATE: DIAGNOSTICS

- Selection Steps Analysis



SELECTION RATE: DIAGNOSTICS

Selection Steps Analysis—Selection Decisions Based on Process (Multiple Hurdles)

- **Hires**

- e.g., Resume review → Exam → Interview)

- **Promotions**

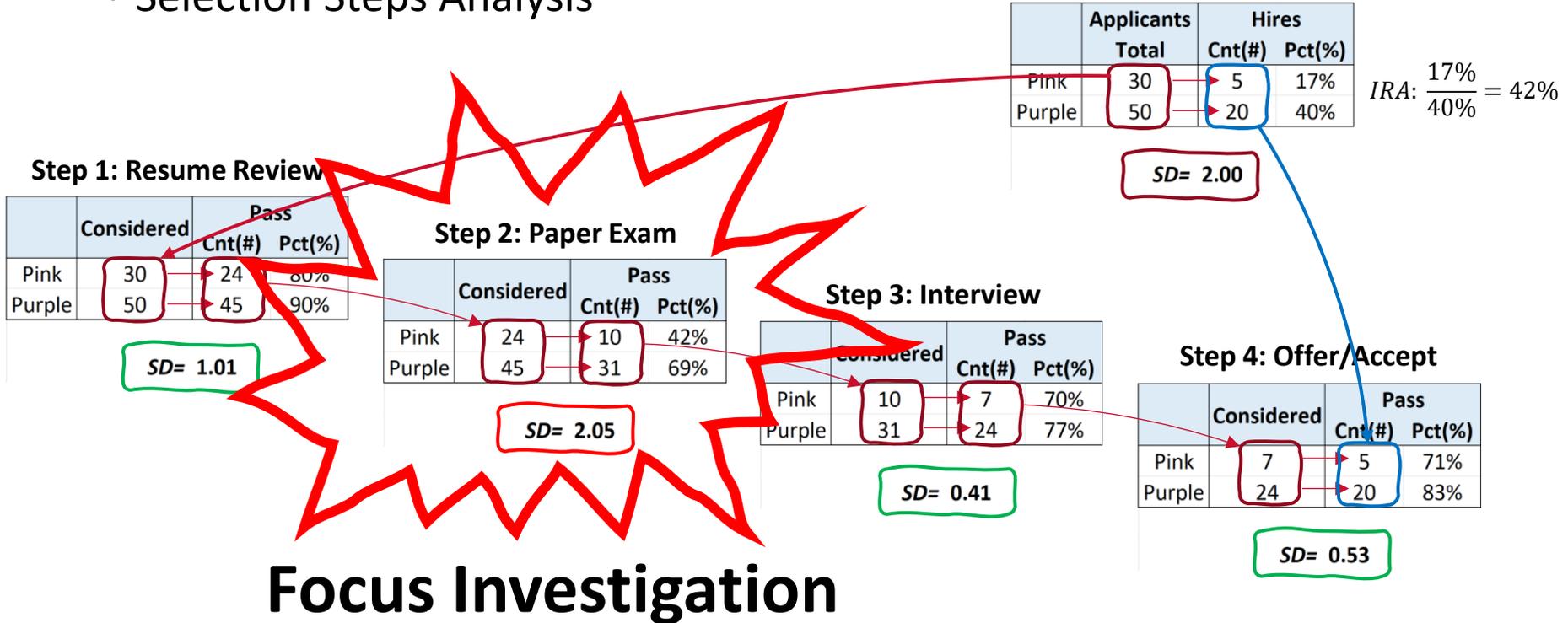
- e.g., Eligibility → Merit-based, Tenure-based, Exam-based → Performance

- **Terminations**

- e.g., Strategy (Skills/Ability, Geography), Decision Matrix

SELECTION RATE: DIAGNOSTICS

- Selection Steps Analysis



SELECTION RATE: DIAGNOSTICS

Are Decisions Job Related and Consistent with Business Necessity?

- **Responding to Statistical Evidence of Potential Adverse Impact**
- **Validity Defense—Are Selection Decisions Valid???**

ARE SELECTION DECISIONS VALID?

UGESP recognizes three types of validation strategies.

- **Criterion-related**: evidence of criterion-related validity should consist of empirical data demonstrating that the selection procedure is predictive of or significantly correlated with important elements of job performance.
- **Content-related**: should consist of data showing that the content of the selection procedure is representative of important aspects of performance on the job for which candidates are to be evaluated.
- **Construct-related**: should consist of data showing that the procedure measures the degree to which candidates have identifiable characteristics which have been determined to be important in successful performance in the job.

RECAP

Inferential Statistics

- Selection Rate Analysis
- Test of Proportions: Binomial
- Multiple Regression: Differences in Means
- Post-Analysis Diagnostic Reviews
 - Uncover reasons for unexplained differences





Thank You

www.nationalilg.org