NATIONAL INDUSTRY LIAISON GROUP

HR Statistics 101
Part 1 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
dkuang@resecon.com

NATIONAL INDUSTRY LIAISON GROUP

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.

NILG NATIONAL INDUSTRY LIAISON GROUP

AGENDA

- HRIS Data
- Data Analysis Basics
  - Get to Know Your Data
  - Data Integrity—cleaning your data
- Data Visualization
- Data Comparisons

HRIS Data

Companies Possess a Wealth of Data from Individuals

Applicant data: demographics, education, qualifications, relevant experience.

Employment outcomes data: placement, job changes, compensation changes, organizational changes.

Performance data: performance reviews, sales, productivity measures.

NILG NATIONAL INDUSTRY LIAISON GROUP

What is Data?

Data could be records in a database, e.g., an **applicant flow** database.

	A	B	C	D	E	F	G	H	I	J	L	M
1	**2 POSITIONS: ASSEMBLER:** Assemblers assist with the overall production of Company products. Some workers are responsible for the disassembly of materials to be refurbished and a											
2	**Operations**											
3	**Applicant**	**Referred by**	**Application**	**Date Applied**	**Resume**	**Phone**	**Phone**	**Testing**	**Test Score**	**Date of**	**Interview**	**Date of**
4		Indeed	Rejected	12/23/2020	link1	1/4/2021	1 - Poor	Pass	87	1/8/2021	2 - Not Good	N/A
5		Indeed	Interview Process	12/27/2020	link2	1/4/2021	3 - Acceptable	Pass	76	1/8/2021	3 - Acceptable	1/25/2021
6		Indeed	Rejected	12/27/2020	link3	1/4/2021	3 - Acceptable	Completed	score not showing	1/10/2021	2 - Not Good	N/A
7		John Doe	Background Chec	12/27/2020	link4	1/4/2021	5 - Excellent	Pass	95	1/11/2021	5 - Excellent	1/25/2021
8		Gerrod Cty Job Fair	Interview Process	12/27/2020	link5	1/4/2021	4 - Worth Conside	Pass	68	1/8/2021	5 - Excellent	1/26/2021
9		Gerrod Cty Job Fair	Interview Process	12/27/2020	link6	1/4/2021	4 - Worth Conside	Completed			4 - Worth Conside	1/27/2021
10		Gerrod Cty Job Fair	Rejected	12/28/2020	link7	1/4/2021	3 - Acceptable	Pass	89	1/8/2021	3 - Acceptable	1/28/2021
11		Gerrod Cty Job Fair	Rejected	12/29/2020	link8	N/A						
12		Gerrod Cty Job Fair	Consider for other	1/5/2021	link9	1/12/2021	4 - Worth Conside	Completed	waiting on score	1/10/2021	4 - Worth Conside	1/29/2021
13		Indeed	Phone Screening	1/6/2021	link10	1/27/2021						
14												
15												

Source: https://fitsmallbusiness.com/applicant-tracking-spreadsheet/

NILG NATIONAL INDUSTRY LIAISON GROUP

What is Data?

Data could be records in a database, e.g., an **employment snapshot**.

ID	*Gender*	*Ethnicity*	*Birth Date*	*Hire Date*	*Job Date*	*Job Group*	*FT/FLSA*	*Grade*	*Job Title*	*Status*	*Annual Rate*
Employee 1	Male	African-American	6/12/1987	11/16/2011	11/15/2011	Job Group 1	Hourly Full Time	Grade 1	Job Title 1	Active	$ 18,200.00
Employee 2	Female	African-American	12/29/1969	9/16/2008	12/14/2010	Job Group 2	Hourly Full Time	Grade 2	Job Title 2	Active	$ 27,768.00
Employee 3	Female	Hispanic	3/16/1958	10/21/2008	9/12/2010	Job Group 1	Hourly Full Time	Grade 2	Job Title 3	Active	$ 20,800.00
Employee 4	Male	Hispanic	3/4/1959	12/2/2008	9/12/2010	Job Group 1	Hourly Full Time	Grade 3	Job Title 4	Active	$ 27,810.00
Employee 5	Female	Caucasian	5/9/1972	10/3/2011	10/2/2011	Job Group 2	Hourly Full Time	Grade 2	Job Title 2	Active	$ 22,256.00
Employee 6	Female	Asian	8/16/1990	1/19/2010	1/19/2010	Job Group 3	Hourly Full Time	Grade 1	Job Title 5	Inactive	$ 18,574.00
Employee 7	Female	African-American	8/31/1980	9/23/2008	9/12/2010	Job Group 2	Hourly Full Time	Grade 2	Job Title 2	Active	$ 23,525.00
Employee 8	Female	Hispanic	8/9/1988	11/4/2008	9/12/2010	Job Group 4	Hourly Full Time	Grade 3	Job Title 6	Active	$ 26,603.00
Employee 9	Female	African-American	11/1/1970	1/26/2010	3/27/2012	Job Group 5	Hourly Full Time	Grade 3	Job Title 7	Active	$ 25,085.00
Employee 10	Female	Hispanic	2/9/1971	9/3/2008	9/12/2010	Job Group 2	Hourly Full Time	Grade 2	Job Title 2	Active	$ 23,733.00
Employee 11	Female	African-American	9/29/1987	9/20/2010	8/29/2011	Job Group 2	Hourly Full Time	Grade 2	Job Title 2	Active	$ 21,840.00
Employee 12	Male	African-American	10/29/1985	10/11/2005	9/12/2010	Job Group 5	Hourly Full Time	Grade 4	Job Title 8	Active	$ 34,486.00
Employee 13	Male	Caucasian	6/21/1966	10/7/2008	12/14/2010	Job Group 2	Hourly Full Time	Grade 2	Job Title 2	Active	$ 27,976.00
Employee 14	Female	Hispanic	8/28/1964	1/26/2010	9/6/2011	Job Group 5	Hourly Full Time	Grade 3	Job Title 7	Active	$ 26,790.00
Employee 15	Female	African-American	4/9/1969	10/7/2008	12/14/2010	Job Group 2	Hourly Full Time	Grade 2	Job Title 2	Active	$ 28,101.00
Employee 16	Female	African American	12/7/1973	5/1/2012	4/30/2012	Job Group 1	Hourly Full Time	Grade 2	Job Title 9	Active	$ 24,398.00
Employee 17	Male	African-American	7/7/1980	6/6/2011	6/30/2011	Job Group 1	Hourly Full Time	Grade 2	Job Title 9	Active	$ 20,800.00

NILG NATIONAL INDUSTRY LIAISON GROUP

Policies and documents are also data

The General Schedule has 15 grades--GS-1 (lowest) to GS-15 (highest). Agencies establish (classify) the grade of each job based on the level of difficulty, responsibility, and qualifications required. Individuals with a high school diploma and no additional experience typically qualify for GS-2 positions; those with a Bachelor's degree for GS-5 positions; and those with a Master's degree for GS-9 positions.

Each grade has 10 step rates (steps 1-10) that are each worth approximately 3 percent of the employee's salary. Within-grade step increases are based on an acceptable level of performance and longevity (waiting periods of 1 year at steps 1-3, 2 years at steps 4-6, and 3 years at steps 7-9). It normally takes 18 years to advance from step 1 to step 10 within a single GS grade if an employee remains in that single grade. However, employees with outstanding (or equivalent) performance ratings may be considered for additional, quality step increases (maximum of one per year).

Source: https://www.opm.gov/policy-data-oversight/pay-leave/pay-systems/general-schedule/

NILG NATIONAL INDUSTRY LIAISON GROUP

Policies and documents are also data

How to Qualify for a WIGI

Eligibility for a within-grade increase is not automatic. It depends on three key factors:

1. **A Permanent Position:** You must be in a permanent position to be eligible for a WIGI.
2. **Acceptable Performance:** Your performance must be at an "Acceptable Level of Competence" (ALOC). This means your work is evaluated as at least fully successful or its equivalent.
3. **Completing the Waiting Period:** You must complete a required waiting period at your current step. The waiting periods get longer as you advance through the steps:
   - **Steps 2, 3, and 4:** 52 weeks (1 year) at the previous step.
   - **Steps 5, 6, and 7:** 104 weeks (2 years) at the previous step.
   - **Steps 8, 9, and 10:** 156 weeks (3 years) at the previous step.

Meeting these requirements allows you to methodically increase your earnings over time, even without a promotion. It is a system that rewards consistent, successful performance and longevity in a role.

Rewarding Excellence: The Quality Step Increase (QSI)

Beyond the standard progression of WGIs, the GS system has a special mechanism to recognize employees who demonstrate outstanding performance. This is the Quality Step Increase, or QSI.

Source: https://www.graduateschool.edu/learn/federal-human-resources/advancing-your-career-in-general-schedule-pay-system

NATIONAL INDUSTRY LIAISON GROUP

Data and Documents are the Foundations of Data Analysis

Statistics blindly run on data may not reflect how decisions were made.

Policies without supporting data cannot be verified.

Robust, informative, and actionable analyses require ***both data and documentation.***

NILG NATIONAL INDUSTRY LIAISON GROUP

Data Analysis Basics

DATA ANALYSIS – BASICS – BEFORE YOU DO ANYTHING

Ensure Buy-in

- Stakeholders
- Data Custodians
- Corporate

Ensure Legal Privilege

- Legal (Internal or External) ***Must*** Be Involved
- Third-party challenges to privilege are common.

DATA ANALYSIS – BASICS

Missing Data

- **MCAR**: Missing Completely At Random
  - No pattern to missing data
- **MAR**: Missing at Random
  - Pattern to missing data but acceptable
- **MNAR**: Missing Not at Random
  - Systemic pattern of missing data that may affect conclusions

DATA ANALYSIS – BASICS

Descriptive Statistics—Numeric Data

- **Numeric Data**: Quantitative, measurable, counts, value, etc.
- **Counts**
- **Range**
  - Minimum
  - Maximum
- **Average**
  - Mean
  - Median
- **Demo/Example**

NILG NATIONAL INDUSTRY LIAISON GROUP

DATA ANALYSIS – BASICS

Descriptive Statistics—Categorical Data

- **Categorical Data**: Qualitative, nominal, labels, groups, etc.
- **Categories**
  - What are the categories?
- **Frequency** (counts) of Categories
- **Cross-Tabulations**
  - Intersectionality of data
- **Demo/Example**

NILG NATIONAL INDUSTRY LIAISON GROUP

Data Visualization

DATA VISUALIZATION – BASICS

Descriptive Statistics—Numeric Data

- **Range**
  - Minimum
  - Maximum
- **Average**
  - Mean
  - Median
- **Demo/Example**

DATA VISUALIZATION – BASICS

Descriptive Statistics—Categorical Data

- **Categories**
  - What are the categories?
- **Frequency** (counts) of Categories
- **Cross-Tabulations**
  - Intersectionality of data
- **Demo/Example**

NILG NATIONAL INDUSTRY LIAISON GROUP

Data Comparisons

Statistics is About Learning from Data, Usually by Making Comparisons

The goal of statistical analyses is to evaluate whether the HR policies and processes are working as intended and in compliance with regulations.

Individuals – *human capital* – are the most important resource of any organization.

HR policies that identify, attract, reward, and retain top human capital foster productivity and a culture to reward merit.

But, Which Comparisons to Make?

Well-defined and monitored HR processes ensure that individuals who possess similar qualifications and interest experience similar outcomes.

Good and complete records and documentation – i.e., good and complete data – allow the analyst to determine the appropriate comparisons and factors to consider.

But Doing Statistics is Hard! We Need Sophisticated Software And Personnel!

Many techniques exist for summarizing data that do not require complex calculations.

These techniques are not just a good first step, they must be the first step in any statistical analysis.

These techniques help identify data errors and gaps in the analyst's understanding of the HR processes.

Selections Comparisons

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}

$$Selection\ Ratio = \frac{\frac{F_{pass}}{F_{pass} + F_{fail}}}{\frac{M_{pass}}{M_{pass} + M_{fail}}}$$

NILG NATIONAL INDUSTRY LIAISON GROUP

Mean or Average Comparisons

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

How to evaluate? Relative to average raises, external benchmarks.

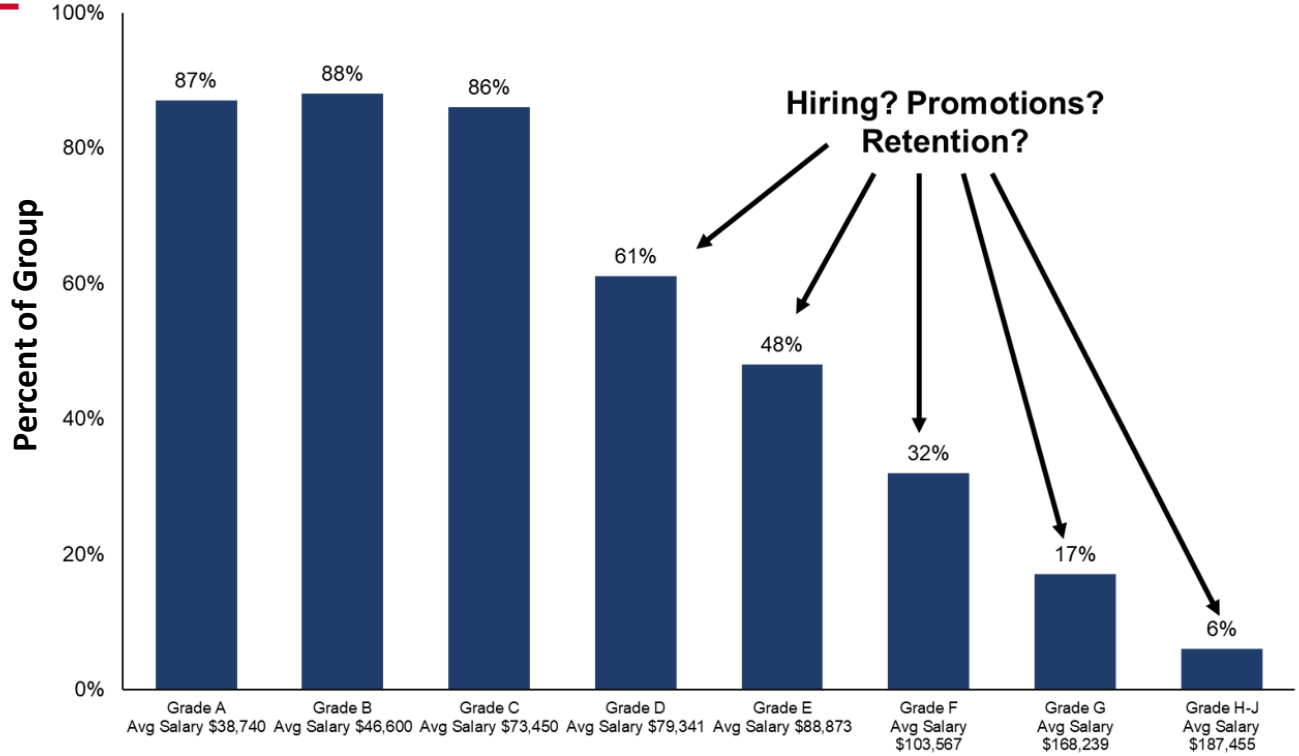
Analysis Group	Female Average Annual Salary	Male Average Annual Salary	Gender Difference	Gender Pay Ratio
Analysis Group 01	$ 71,112	$ 66,459	-$ 5,347	107.0%
Analysis Group 02	$ 96,588	$ 96,906	-$ 318	99.7%
…				
Analysis Group 10	$ 151,462	$ 149,155	$ 2,307	101.5%
Analysis Group 11	$ 161,605	$ 167,405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

When Differences Appear to Exist: Cohort/Group Reviews

ID	Gender	Ethnicity	Hire Date	Job Date	Analysis Group	FT/FLSA	Grade	Job Title	Status	Annual Rate
1222	Male	White	05/06/21	05/06/21	Analysis Group 1	Hourly Full Time	Grade 2	Job Title 3	Active	$ 60,030.00
723	Male	Hispanic	07/01/21	07/01/21	Analysis Group 1	Hourly Full Time	Grade 2	Job Title 3	Active	$ 60,176.45
805	Female	African-American	10/30/14	02/22/20	Analysis Group 1	Hourly Full Time	Grade 2	Job Title 3	Active	$ 61,048.97
722	Male	African-American	01/03/17	09/12/21	Analysis Group 1	Hourly Full Time	Grade 2	Job Title 3	Active	$ 61,439.84
635	Female	Hispanic	09/10/19	09/10/19	Analysis Group 1	Hourly Full Time	Grade 2	Job Title 3	Active	$ 62,400.00
1107	Female	White	04/06/14	09/10/19	Analysis Group 1	Hourly Full Time	Grade 2	Job Title 3	Active	$ 62,400.00
6053	Female	White	02/08/15	12/11/19	Analysis Group 1	Hourly Full Time	Grade 2	Job Title 3	Active	$ 62,400.00
1088	Male	African-American	10/15/16	10/13/20	Analysis Group 1	Hourly Full Time	Grade 2	Job Title 3	Active	$ 72,759.00
744	Female	White	06/02/19	06/02/19	Analysis Group 1	Hourly Full Time	Grade 2	Job Title 3	Active	$ 74,656.01
945	Male	White	09/10/21	09/10/21	Analysis Group 1	Hourly Full Time	Grade 2	Job Title 3	Active	$ 75,753.00
575	Female	White	02/14/22	02/14/22	Analysis Group 11	Sal Exempt Full Time	Grade 9	Job Title 24	Active	$ 139,932.00
1438	Male	Hispanic	05/16/21	05/16/21	Analysis Group 11	Sal Exempt Full Time	Grade 9	Job Title 24	Active	$ 142,906.88
842	Female	White	02/05/21	02/05/21	Analysis Group 11	Sal Exempt Full Time	Grade 9	Job Title 24	Active	$ 169,505.04
735	Female	Asian	04/25/15	04/25/15	Analysis Group 11	Sal Exempt Full Time	Grade 9	Job Title 24	Active	$ 172,115.68

NILG NATIONAL INDUSTRY LIAISON GROUP

Always Visualize Your Data!



NILG NATIONAL INDUSTRY LIAISON GROUP

When simple comparisons are not enough.

The average temperature in Dallas, TX and Guatemala City, Guatemala is 68 degrees.

Where will a 10-degree change in temperature be more surprising?

In Dallas, TX, where monthly average temperatures range between 47 and 87 degrees?

In Guatemala City, Guatemala, where monthly average temperatures range between 62 and 71 degrees?

Statistical significance helps quantify deviations from a baseline.

So, What's Statistical Significance?

Statistical significance is a measure of ***surprise***.

Is the observed outcome ***extreme or surprising***, ***given*** the assumed baseline?

For example: If, on average, members of two groups are equally skilled and meet the requirements of the position?

Is a 3% difference in hiring or promotion rates surprising?

Is a $5,000 difference in base or incentive pay surprising?

NILG NATIONAL INDUSTRY LIAISON GROUP

So, What's Statistical Significance?

In many academic fields and legal matters, statistical significance is defined as outcome differences equivalent to 2 more standard deviations (SDs).

Why 2+ SDs? Why not 1 SD or 5 SDs?

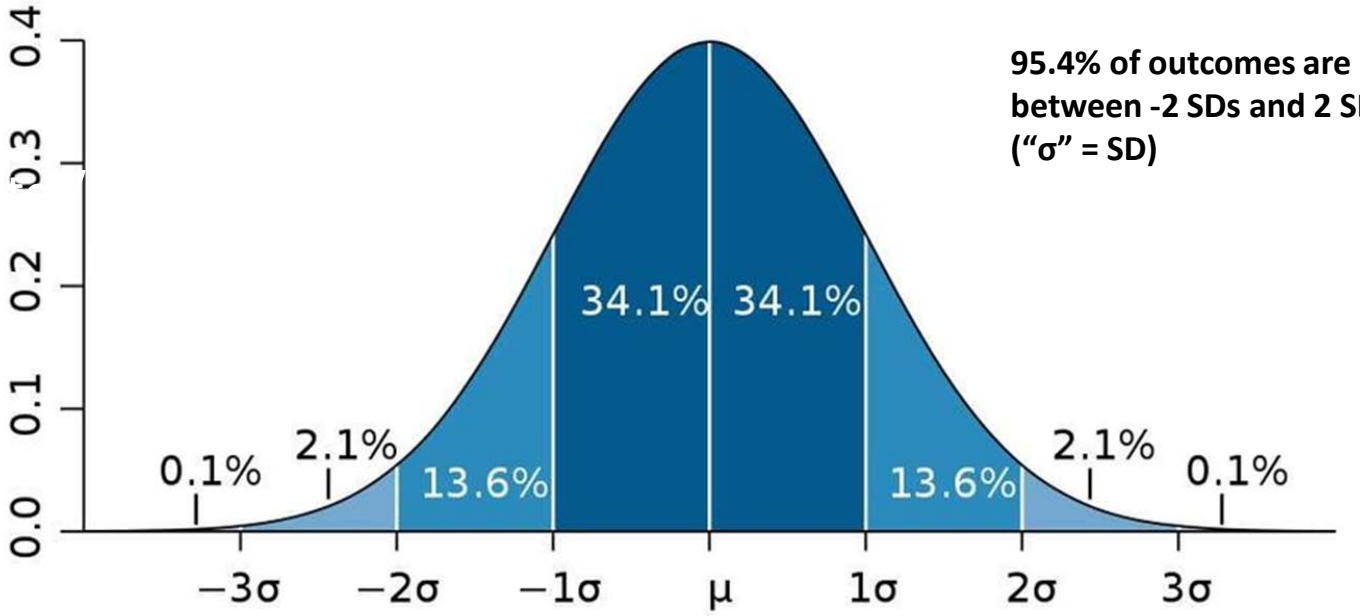
Outcomes equivalent to 2+ standard deviations are expected to occur less than 5% of the time.

In several scientific fields, events expected to happen less than 5% of the time are considered rare enough to require additional review.

As statistics became prevalent in litigation, the courts adopted the "2+ SDs" standard.

NILG NATIONAL INDUSTRY LIAISON GROUP

The Bell Curve: 2+ SDs ≈ Less than 5% Chance



95.4% of outcomes are between -2 SDs and 2 SDs ("σ" = SD)

Source: https://news.mit.edu/2012/explained-sigma-0209

NILG NATIONAL INDUSTRY LIAISON GROUP

Interpreting “Statistical Significance”

Statistical significance does not mean a difference exists.

Statistical significance means that further research is needed.

A correlation between outcomes and group membership exists.

Statistical significance does not prove discrimination.

The only inference an analyst can make is that there is a difference correlated with group membership.

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

Interpreting “Statistical Significance”

Statistical inference is more than checking whether the SDs of a difference are above or below 2.

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

What can you do to minimize these errors?

Good data: Data errors cause both Type I and Type II errors

Good documentation: Incorrect comparisons cause both Type 1 and Type II errors

NILG NATIONAL INDUSTRY LIAISON GROUP

Statistical Significance is Not Practical Significance

A statistically significant difference may lack practical relevance.

Large databases act as magnifying glasses, allowing the analyst to detect smaller differences.

> According to the Current Population Survey, the US mean annual wage is $67,920.
>
> Given the size of the survey, a $136 difference in mean wages would be statistically significant using the 2-SD threshold.

HR databases continue to grow, increasing the chances of finding statistically significant yet immaterial differences.

NILG NATIONAL INDUSTRY LIAISON GROUP

Mean or Average Comparisons with SDs

Analysis Group	Females v. Males	
	Percentage Difference	Standard Deviations
Analysis Group 01	7.0%	1.87
Analysis Group 02	-0.3%	-0.86
…		
Analysis Group 20	0.6%	1.71
Analysis Group 21	- 4.8%	- 2.73
Analysis Group 22	3.9%	2.63

A statistically significant difference indicates that the group needs additional review.

NILG NATIONAL INDUSTRY LIAISON GROUP

Where do we start?

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

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, ***always look at your data.***	

NILG NATIONAL INDUSTRY LIAISON GROUP

RECAP

- HRIS Data
- Data Analysis Basics
  - Get to Know Your Data
  - Data Integrity—cleaning your data
- Data Visualization
- Data Comparisons

PREVIEW HR STATS 101 – PART 2

Inferential Statistics

- Selection Rate Analysis
- Test of Proportions: Binomial
- Multiple Regression: Differences in Means
- Diagnostic Reviews (i.e., deep dives to uncover reasons for unexplained differences)

NATIONAL INDUSTRY LIAISON GROUP

Thank You

www.nationalilg.org