People Analytics

Groupwork Use Case 1 – Diversity and Analytics

In this workshop, you will use R to explore diversity issues in an organization through descriptive, inferential, and predictive statistical techniques. You will work with gender, ethnicity, and organizational data across teams and functions.

Prompt 1: Gender and Job Grade (Chi-Square Test)

The organization 'SlidesRUs' has shared its gender composition across job grades. You are tasked with determining whether gender is associated with job grade using a chi-square test.

Tasks:

- Import 'Ch4_diversity.xlsx' into R.
- Recode the gender variable into 'Men' and 'Women'.
- Generate frequency and cross-tabulation tables.
- Run a chi-square test to determine the statistical significance of the observed patterns.

Deliverables:

- Table of observed vs. expected frequencies.
- Chi-square output with interpretation.
- Brief reflection: What might explain the disparity in gender across grades?

Prompt 2: Descriptive Statistics on Ethnic Diversity Across Teams

You now have access to the dataset 'Ch4_div_group.xlsx' from a financial organization. It contains team-level data, including gender and UG (underrepresented group) statistics.

Tasks:

- Import the dataset into R.
- Generate descriptive statistics for PercentMale and UG.
- Interpret means, standard deviations, minimum and maximum values.

Deliverables:

- Table with descriptive statistics.
- Short analysis: What do the results suggest about gender and UG diversity across teams?

Prompt 3: Comparing UG and Gender Across Functions (t-test)

You want to test whether the sales and professional service functions differ in terms of UG representation and gender ratio.

Tasks:

- Recode 'Function' variable into 'Sales' and 'Professional Service'.
- Conduct independent t-tests on PercentMale and UG.
- Run Levene's Test to check for equal variances.

Deliverables:

- t-test output and interpretation for UG and PercentMale.
- Levene's test result and its implication.
- Summary: Do the differences point to potential diversity challenges?

Prompt 4: Predicting UG with Multiple Regression

Using multiple regression, explore which organizational variables predict the proportion of UG in a team.

Tasks:

- Run a regression model with UG as the dependent variable.
- Include predictors: LondonorNot, Function, GroupSize, NumberFeMaleTeamLeads, PercentMale.
- Interpret the coefficients and R-squared value.
- Identify which predictors significantly affect UG levels.

Deliverables:

- Regression output with interpretation.
- Summary table of standardized coefficients.
- Executive summary: Based on this model, what would you recommend the company do next?