

EDGE Certified Foundation



EDGEplus UNEXPLAINED GENDER AND INTERSECTIONAL PAY GAP (UGIPG) ANALYSIS

METHODOLOGY V1.3
1 NOVEMBER 2024





CONTENTS

1. PURPOSE
 - 1.1. Scope of methodology
2. DATA REQUIREMENTS
 - 2.1. Selection of a reference period
 - 2.2. Employees included in the analysis
 - 2.3. Dependent variables included in the analysis
 - 2.4. Explanatory variables included in the analysis
3. EDGEPLUS-COMPLIANT ANALYSIS METHODOLOGY
 - 3.1. Linear regression analysis
 - 3.2. Reported results of the regression analysis



1. PURPOSE

The purpose of this document is to describe in detail the method prescribed by the EDGE Certified Foundation to determine whether there are unexplained pay gaps between women and men of different races/ethnicities.

This method extends the EDGE-compliant Unexplained Gender Pay Gap (UGPG) analysis that is generally applied as part of the preparation for EDGE Certification, in order to enable a gender and intersectional analysis of pay gaps within certifying organizations. The EDGE-compliant UGPG analysis may not include discriminatory variables such as race/ethnicity. However, these factors can be important drivers of inequalities in pay between women and men. This method describes the approach applied to measure the importance of such factors in explaining pay.

Throughout this document, the method prescribed by the EDGE Certified Foundation is referred to as an EDGEplus-compliant analysis. For further information concerning the method, please contact:

Email: aunguresan@edge-cert.org

Website: www.edge-cert.org

Mailing address: Ms Aniela Unguresan
Founder
EDGE Certified Foundation
Vorstadt 2, 6300 Zug
Switzerland

1.1. Scope of Methodology

The method prescribed by the EDGE Certified Foundation has been developed for the purposes of measuring the intersectionality between gender and race/ethnicity – as part of the EDGEplus analysis. An organization is encouraged to also apply the methodology independently, to determine whether there are any unexplained pay gaps between women and men of different races/ethnicities. The method described in this document shall be used exclusively for the purposes of analysing intersectional pay equity issues as part of the EDGEplus analysis. To determine whether there are unexplained pay gaps between women and men for the purpose of achieving EDGE Certification, an organization must apply the EDGE-complaint analysis described in the “EDGE Unexplained Gender Pay Gap (UGPG) Analysis Methodology”.

2. DATA REQUIREMENTS

2.1. Selection of a reference period

To perform an EDGEplus-compliant UGIPG analysis, the organization chooses the most recent 12-month reference period against which the assessment will be conducted. The reference period may be a calendar year, fiscal year or some other period that is relevant to the organization for internal reporting purposes. The first reference period becomes the baseline for assessing an organization’s progress. Therefore, the months chosen for the reference period should not change in subsequent reference periods unless there are exceptional circumstances like a merger or acquisition that results in a change to an organization’s reporting period, or in the case of a global approach, the alignment of multiple country-entity reference periods.

To be considered EDGEplus-compliant, no more than 12 months should have passed since the end of the reference period used for the EDGEplus-compliant analysis.



2.2. Employees included in the analysis

Anyone on the payroll of the organization (e.g., the country-level entity) undergoing the analysis should be included in the analysis, including employees of acquired organizations, where the acquired organization has been fully integrated on a legal basis into the acquirer, who are employed on the same policies and practices as the organization. The minimum number of observations (i.e. number of employees to be included in the analysis) required to run an EDGEplus-compliant regression analysis is addressed in section 2.5.

2.3. Dependent variables included in the analysis

The regression analysis shall be run as a set of two regression analyses. The first shall consider 'Salary' (base salary) as the dependent variable and the second shall consider 'Pay' (base salary plus bonuses and other cash benefits) as the dependent variable. Inclusion and exclusion criteria for Salary and Pay are shown in Table 1 below.

As a general principle, if applicable, the organizations shall follow regulatory/legislative criteria for inclusion/exclusion and categorization of base pay and bonuses/benefits for relevant jurisdiction.

1. Salary (base salary) and Pay (base salary plus bonuses and other cash benefits) shall be actual (realized) remuneration – e.g., remuneration that has actually been paid to employees during the Reference Period, not target salary or pay (bonuses/benefits) not yet paid.
2. Salary (base salary) and Pay (base salary plus bonuses and other cash benefits) paid during the Reference Period for specified periods of time of less than or in excess of a year shall be pro-rated to annualized amounts based on contractual hours.

For example, person A was employed on 1st March at 100% with an annual base salary of 60,000. Her annualized salary for the full calendar year is 60,000.

3. Salary (base salary) and Pay (base salary plus bonuses and other cash benefits) paid during the Reference Period for less than full-time work shall be pro-rated to Full-Time Equivalent amounts based on contractual hours.

For example, person B is employed part time, at 70%, with an annual base salary of 68,000. His annual base salary pro-rated to 100% is

$$\frac{68,000}{7} * 10 = 97,143.$$



Table 1. Salary and Pay inclusion and exclusion criteria

| INCLUDED | | EXCLUDED |
|--|--|--|
| SALARY | PAY | |
| <ul style="list-style-type: none"> ✓ In principle, all that is considered "standard" or "ordinary" for that kind of job or position | <ul style="list-style-type: none"> ✓ In principle, Base Salary plus all that is negotiable, variable and/or on top of the Base Salary; including recurring payments | <ul style="list-style-type: none"> ✗ In principle, one-off incentives or bonuses, e.g., signing bonus, "golden hello", or lump sum payment at beginning of a work contract to compensate for lost potential bonus at previous employer; redundancy or termination payments |
| <ul style="list-style-type: none"> ✓ Gross salary specified in employment contract, e.g., before tax, before other deductions (e.g., employee pension contributions, salary sacrifice/salary packages) | <ul style="list-style-type: none"> ✓ If unable to distinguish between "bonus" payments and "overtime" and/or "shift" pay/premiums, then include overtime and/or shift pay/premiums in bonus calculation | <ul style="list-style-type: none"> ✗ Overtime or shift pay/premiums ✗ Expense reimbursements ✗ Non-'cash' benefits (e.g., company cars; health (e.g., private) insurance) |
| <p>Extras/Allowances that are considered standard for the type of job/position:</p> <ul style="list-style-type: none"> ✓ On-call allowances/extras (e.g., for emergency workers) ✓ Allowances and/or awards that are non-negotiable and applied uniformly to employees for that kind of position | <p>Bonuses based on performance (organization's and/or employee's):</p> <ul style="list-style-type: none"> ✓ Incentives ✓ Commissions ✓ Awards with a monetary value (e.g., vouchers, securities & securities options/interests – shares, bonds, debentures and futures) ✓ Full amount of bonuses awarded in a particular year but payable over multiple years | <ul style="list-style-type: none"> ✗ Deferred bonus plans (e.g., not incurring a charge to income tax during the Reference Period) ✗ Allowances that are applied to ALL eligible employees AND are non-negotiable AND are non-job/non-position related (e.g., per child allowances; location-specific weightings/allowances) |
| <ul style="list-style-type: none"> ✓ Maternity/Parental/Paternity pay ✓ Sick Pay ✓ Holiday Pay ✓ Other special leave pay (e.g., long-service leave; carer's leave; etc.) | <ul style="list-style-type: none"> ✓ Extras/Allowances that are negotiable | <ul style="list-style-type: none"> ✗ Pay in lieu of annual or other leave |



2.4. Explanatory variables included in the analysis

There are two possible levels of complexity permitted when it comes to performing EDGEplus-compliant regression analyses:

1. Assessing potential unexplained gender pay gaps using an **EDGEplus Standard Regression Analysis**. An **EDGEplus Standard Regression Analysis** is an EDGEplus-compliant regression analysis that measures salary and pay against the common set of variables described below under **Standard Variables**.
2. Assessing potential unexplained gender pay gaps using an **EDGEplus Customized Regression Analysis**. An **EDGEplus Customized Regression Analysis** is an EDGEplus-compliant regression analysis that measures salary and pay against the common set of **EDGEplus Standard Variables** and additional organization-specific variables – which shall only be permitted if an organization performs the EDGEplus Standard Regression Analysis first or in conjunction with the customized analysis.

The organization shall use the same Regression Analysis method for the analyses of both ‘Salary’ and ‘Pay’ as dependent variables (i.e., Standard or Customized Regression).

Regression analysis calculates the extent the variations in the dependent variable (pay or salary), can be explained by variations in the explanatory variables (gender and race/ethnicity), when the control variables (a collection of different characteristics related to the employees and the workplace) are kept constant. Explanatory and control variables are called independent variables, for statistical and computational purposes, independent variables are transformed into ‘predictors’ when writing the code for the analysis.

The total number of predictors will depend on whether a variable is considered:

- a) ‘Numerical’ – a real number, e.g., age or tenure. Each numerical variable counts as one predictor.
- b) ‘Categorical’ – a variable that uses two or more possible values, usually of a qualitative nature, e.g., level of responsibility or department:
 - i. Where a categorical variable has only two possible values, it is coded as a single binary variable and is counted as one predictor. For instance, the gender variable can be coded as a binary variable, equal to 0 for male and equal to 1 for female.
 - ii. Where a categorical variable has more than two possible values (categories), it is transformed into a collection of binary variables, and the number of additional predictors will be the number of categories minus one (K-1).

Standard Variables

An EDGEplus Standard Regression Analysis shall measure salary and pay (as defined above) against the following pre-defined variables:

1. **Gender** – the first explanatory variable, coded as binary (0 for Male, 1 for Female), which counts as one predictor.
2. **Tenure** – a numerical variable usually inputted as years of service in the organization, which counts as one predictor.
3. **Age** – a numerical variable, which counts as one predictor.



4. *Age²* – age squared – a numerical variable included for statistical purposes, which counts as one predictor.
5. *Responsibility of the role* – a categorical variable coded as either Individual Contributor (IC) or People Manager (PM), which counts as one predictor.
 - i. An Individual Contributor is an employee who is a team member; frontline staff member; a solo operator like a consultant, lawyer, researcher, literary agent, software engineer; and so on: an employee who is not responsible for managing or supervising any other employees.
 - ii. A People Manager is an employee in a supervisory role, responsible for managing one or more people, and could include being a line manager; performance manager; project or team leader; section, division, or department head; and so on.
6. *Type of performed function* – a categorical variable coded as either Support Function (SF) or Core Function (CF), which counts as one predictor.
 - i. A Support Function is a supporting activity carried out by the organization in order to enable or to facilitate the core business functions (e. g. legal services, accounting, book-keeping and auditing, human resources management, payroll management, procurement functions). The outputs (results) of support business functions are not themselves intended directly for the market or for third parties.
 - ii. A Core Function is defined as an activity of an organization yielding income: the production of final goods or services intended for the market or for third parties. Usually, the core business functions make up the primary activity of the organization (e.g., client facing roles, production).
7. *Level of responsibility* – a categorical variable with five levels, which counts as four predictors. The organization may choose to code this variable with four levels, in which case it counts as three predictors.
 - i. Top management – reporting formally and directly to the CEO.
 - ii. Upper management – reporting formally and directly to the top management.
 - iii. Middle management – reporting formally and directly to the upper management.
 - iv. Junior management – reporting formally and directly to the middle management.
 - v. Operational level – all other employees.

In addition, an EDGEplus Standard Regression Analysis shall measure salary and pay (as defined above) against the following additional explanatory variable:

1. *Race/Ethnicity* – a categorical variable with at least two and up to twenty categories, coded as a series of at least one and up to nineteen binary predictors, where the number and labels of the categories are to be determined by the organization and where '0' corresponds to a chosen baseline category. For example, the six categories: 'White', 'Asian', 'Black', 'Hispanic', 'Indigenous' and 'Other' with 'White' as the baseline, which would be coded as a series of five binary predictors.

As a general principle, if applicable, an organization must follow regulatory/legislative criteria for the inclusion and categorization of explanatory variables for race/ ethnicity for relevant jurisdiction.



Customized Variables

An EDGEplus Customized Regression Analysis shall measure salary and pay (as defined above) against the standard variables plus additional variables defined by the organization in line with the organization's compensation and benefits structure. For example, but not limited to:

1. *Level of responsibility* may be separated into more levels (and therefore more predictors) than the EDGE five levels of responsibility.
2. *Education / qualifications*
3. *Business unit*
4. *Department*
5. *Geographic area*

An EDGEplus Customized Regression Analysis shall not include any discriminatory variables, such as part-time or any other variable that is not strictly related to the skills and/or competencies of the employee, or the nature of the job.

2.5. Minimum number of observations (employees) required to run the analysis

Dependent on whether four or five levels of responsibility are included and the number of categories chosen for the explanatory variable race/ethnicity, an EDGEplus Standard Regression Analysis shall have between 11 and 48 of the above pre-defined standard predictors, including the explanatory predictor of gender and the predictors of the interaction term between gender and race/ethnicity (see section 3.1). The minimum number of standard predictors (i.e. 11) corresponds to an EDGEplus Standard Regression with four levels of responsibility and only two levels for the explanatory variable race/ethnicity. The maximum number of standard predictors (i.e. 48) corresponds to an EDGEplus Standard Regression with five levels of responsibility and twenty levels for the explanatory variable race/ethnicity. Given at least ten observations per predictor are required to ensure there is sufficient statistical power to achieve credible results, an EDGEplus Standard Regression Analysis requires a minimum number of observations that varies from 110 to 480, i.e. uses relevant data on at least 110 to 480 employees, depending on the number of categories set by the user for the explanatory variable race/ethnicity. Consistently, the data requirements of an EDGEplus Customized Regression Analysis increase with the number of custom predictors added to the standard set of variables.

3. EDGEPLUS-COMPLIANT ANALYSIS METHODOLOGY

An EDGEplus-compliant analysis shall be performed using Ordinary Least Squares (OLS) linear regression. Linear regression is a statistical method that can quantitatively demonstrate whether an organization has any pay gap which cannot be explained by any other factors than gender and race/ethnicity, e.g., unexplained gender pay gaps. The following section sets out the methodology behind this analysis.

3.1. Linear regression analysis

Linear regression is a statistical method that mathematically explains a chosen dependent variable (e.g., an employee's wage) in terms of a selection of independent predictors (e.g., gender, age, level of



responsibility, etc.). Use of this method allows an organization to assess the extent to which an employee's salary or pay is influenced by gender after accounting for all other factors.

More specifically, linear regression expresses the dependent variable as a linear function of the independent predictors under certain assumptions about their statistical distribution. The general form of the regression equation is

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi} + \varepsilon_i$$

where:

- y_i is the dependent variable for a given employee i , e.g., salary or pay.
- β_0 is a constant representing the average of the response y_i when all independent predictors are set to 0.
- $\beta_1 \dots \beta_p$ are the linear regression coefficients associated with the 1 ... p independent predictors.
- $x_{1i} \dots x_{pi}$ are the values of the p independent predictors for a given employee i , e.g., the gender, age, level of responsibility, etc. for a given employee.
- ε_i is the error term of a given employee i , the difference between the predicted dependent variable and its true value (for example, between an employee's actual salary and their salary according to the linear regression equation).

The OLS approach calculates the values of the coefficients $\beta_1 \dots \beta_p$ required to minimize the error terms ε_i .

By making the statistical assumption that the distribution of the error terms ε_i is normal with zero mean and constant variance, p-values and confidence intervals can also be used to indicate the level of uncertainty associated with the coefficient estimates.

EDGEplus Standard Regression Analysis

The form of the regression equation adopted in the EDGEplus-compliant Standard Regression Analysis for salary shall be

$$\ln(\text{Salary}_i) = \beta_0 + \beta_1 \text{Gender}_i + \beta_2 \text{Race/Ethnicity}_i + \beta_3 \text{Gender}_i * \text{Race/Ethnicity}_i + \beta_4 \text{Tenure}_i + \beta_5 \text{Age}_i + \beta_6 \text{Age}_i^2 + \beta_7 \text{Responsibility of Role}_i + \beta_8 \text{Type of Performed Function}_i + \beta_9 \text{Level of Responsibility}_i + \varepsilon_i$$

where:

- $\ln(\text{Salary}_i)$ is the log-transformed salary of a given employee i .
- Gender_i , Tenure_i , Age_i etc. are the values of the standard predictors described in section 2.4 above for a given employee i .
- Race/Ethnicity_i is the value of the standard predictor for race/ethnicity described in section 2.4 above for a given employee i .
- $\text{Gender}_i * \text{Race/Ethnicity}_i$ is the value of the standard predictor for a statistical term called interaction between gender and race/ethnicity, which highlights that race/ethnicity moderates the effect of gender on salary/pay. In other words, for each race/ethnicity category, the effect of gender on salary/pay is different from other race/ethnicity categories.

Similarly, the form of the regression equation adopted in the EDGEplus-compliant Standard Regression Analysis for pay is



$$\ln(\text{Pay}_i) = \beta_0 + \beta_1 \text{Gender}_i + \beta_2 \text{Race/Ethnicity}_i + \beta_3 \text{Gender}_i * \text{Race/Ethnicity}_i + \beta_4 \text{Tenure}_i + \beta_5 \text{Age}_i \\ + \beta_6 \text{Age}_i^2 + \beta_7 \text{Responsibility of Role}_i + \beta_8 \text{Type of Performed Function}_i \\ + \beta_9 \text{Level of Responsibility}_i + \varepsilon_i$$

where $\ln(\text{Pay}_i)$ is the log-transformed pay of a given employee i .

EDGEplus Customized Regression Analysis

The form of the regression equation adopted in the EDGEplus-compliant Customized Regression Analysis shall be as in the standard regression analysis above, but with the addition of predictors defined by the organization in line with the organization's compensation and benefits structure.

3.2. Reported results of the regression analysis

An EDGEplus-compliant regression analysis shall report several metrics and statistics to inform the assessment of wage differences between men and women. These metrics and statistics are detailed in the following section.

Regression Coefficients

Estimated values for the coefficients of the EDGEplus Standard and (if performed) EDGEplus Customized Regression analyses shall be reported. These values are used to calculate the **Effect of gender or EDGEplus UGPG** as described below.

Adjusted R-Squared

The Adjusted R-Squared value is an output of most statistical software packages and is a measure of model fit. More specifically, the Adjusted R-Squared value measures the amount of variation in the data that is explained by the linear regression model. Values close to one indicate that variation in employees' salary and pay can be almost entirely explained by the predictors included in the EDGEplus Standard or EDGEplus Customized Regression analysis. On the other hand, values close to zero indicate that variations in salary and pay cannot be adequately explained by the regression analysis and are likely to be driven by factors outside of the analysis.

From an EDGEplus compliance perspective, when an organization performs an EDGEplus Customized Regression after or in conjunction with an EDGEplus Standard Regression, the adjusted R-squared value for the EDGEplus Customized Regression *shall not* be lower than that produced by the EDGEplus Standard Regression. If the inclusion of customized predictors lowers the adjusted R-squared value, then a different subset of customized predictors shall be selected and the regression re-run. If no subset can be found that leads to a higher adjusted R-squared value, the organization shall exclude the results and use only the results of the EDGEplus Standard Regression.

Observations

The number of observations included in the EDGEplus Standard or Customized Regression Analysis shall be reported. This is the number of employees whose data were input for analysis and should correspond to the total numbers of employees in the organization as defined in section 2.5.

Number of predictors

The number of predictors in the EDGEplus Standard or Customized Regression Analysis shall be reported. This value refers to the number of predictors used in the analysis, which corresponds to the



number of variables that were included in the regression model, taking into account the type of variables. Categorical variables are treated as multiple predictors for computational purposes, the total number of which depends on the allowed values.

Effect of gender or EDGEplus UGPG

To perform an EDGEplus-compliant analysis, an organization should report the EDGEplus UGPG for each combination of the categories for gender, race/ethnicity in comparison to every other combination.

To estimate the EDGEplus UGPG, the estimated coefficients β_1 , β_2 , β_3 for the predictors corresponding to each combination of the categories for gender, race/ethnicity and their interactions should be transformed using the following equation:

$$\left(\frac{\exp(\beta_{1i} + \beta_{2i} + \beta_{3i})}{\exp(\beta_{1j} + \beta_{2j} + \beta_{3j})} - 1 \right) * 100\%,$$

where β_{1i} , β_{2i} , β_{3i} are the estimated coefficients for gender, race/ethnicity and their interaction corresponding to category i , and β_{1j} , β_{2j} , β_{3j} are the estimated coefficients corresponding to any baseline category j of choice. Note that this is the chosen baseline category to compare two categories, and not necessarily the baseline category that the EDGE Empower[®] Pay Tool is using for statistical analysis (statistical baseline). It means that these coefficients highlight the impact of moving from the statistical baseline category to any chosen category. Therefore, by default, the coefficients of the statistical baseline category are equal to zero.

This formula expresses the effect of the intersection of gender and race/ethnicity on salary or pay as a percentage relative to the chosen baseline category. This formula should be applied to all combinations of the estimated coefficients for gender and race/ethnicity (for example, Male and Black vs. Male and White, Female and Black vs. Male and White, Female and Black vs. Female and White, etc.). The values can be interpreted as the average percent increase (or decrease) in salary or pay for a given combination of gender with race/ethnicity in comparison to a baseline category, where:

- a) A positive (+) value denotes a pay gap in favour of the given combination of gender and race/ethnicity.
- b) A negative (-) value denotes a pay gap in favour of the baseline category.

For example, suppose an organization chooses to run the Regression Analysis with the explanatory variable race, where race is either 'White', 'Black' or 'Other' and 'White' is the chosen baseline category. Suppose that the coefficient for gender is found to be -0.15, the coefficient for race 'Black' is found to be 0.01, the coefficient for race 'Other' is found to be -0.01, the coefficient for the interaction between gender and race 'Female * Black' is found to be -0.05 and, finally, the coefficient for the interaction between gender and race 'Female * Other' is found to be -0.06. By default, the coefficients corresponding to 'Male' and 'White' are 0 (the statistical baseline category). The EDGEplus UGPG for Black women (category i) in comparison to White men (category j) is

$$\left(\frac{\exp(-0.15 + 0.01 - 0.05)}{\exp(0 + 0 + 0)} - 1 \right) * 100\% = -17.3\%.$$



This indicates that when all other personal and organization-related factors are accounted for, Black women are paid on average 17.3% less than the baseline category, White men. Similarly, the EDGEplus UGPG for Black women in comparison to Black men is

$$\left(\frac{\exp(-0.15 + 0.01 - 0.05)}{\exp(0 + 0.01 + 0)} - 1 \right) * 100\% = -18.1\%.$$

This indicates that when all other personal and organization-related factors are accounted for, Black women are paid on average 18.3% less than Black men. Completing the example above, the organization would then report the following values:

| EDGEplus UGPG | | Baseline category (<i>j</i>) | | | | | |
|----------------------------------|-------------------|--------------------------------|-----------------|-----------------|-------------------|-------------------|-------------------|
| | | Male * White | Male * Black | Male * Other | Female * White | Female * Black | Female * Other |
| Comparison category (<i>i</i>) | Male * White | | -1.00% | 1.01% | 16.18% | 20.92% | 24.61% |
| | Male * Black | 1.01% | | 2.02% | 17.35% | 22.14% | 25.86% |
| | Male * Other | -1.00% | -1.98% | | 15.03% | 19.72% | 23.37% |
| | Female * White | -13.93% | -14.79% | -13.06% | | 4.08% | 7.25% |
| | Female * Black | -17.30% | -18.13% | -16.47% | -3.92% | | 3.05% |
| | Female * Other | -19.75% | -20.55% | -18.94% | -6.76% | -2.96% | |