

**Figure 1:** Normality Assessment of Pollutant Levels and Affected Areas: Results from Kolmogorov-Smirnov and Shapiro-Wilk Tests

**Normality Test Results**

This table presents the outcomes of the Kolmogorov-Smirnov and Shapiro-Wilk tests for assessing the normality of two variables: **"Pollutant\_Level"** and **"Count\_of\_Affected\_Areas."**

**1. Significance Levels**

* **p-values**: Both variables have p-values **less than 0.05** in both tests.
* **Interpretation**: This indicates that neither variable follows a normal distribution.

**2. Normality Insight**

* A p-value below 0.05 suggests a **significant deviation from normality**, meaning the data distribution for these variables does not meet the normality assumption required for parametric tests.

**Test Descriptions**

**Kolmogorov-Smirnov Test**

* **Purpose**: Nonparametric test comparing the sample data distribution to a theoretical distribution (e.g., normal distribution).
* **Lilliefors Correction**: Applied to improve accuracy when population parameters (mean and standard deviation) are estimated from the sample.

**Shapiro-Wilk Test**

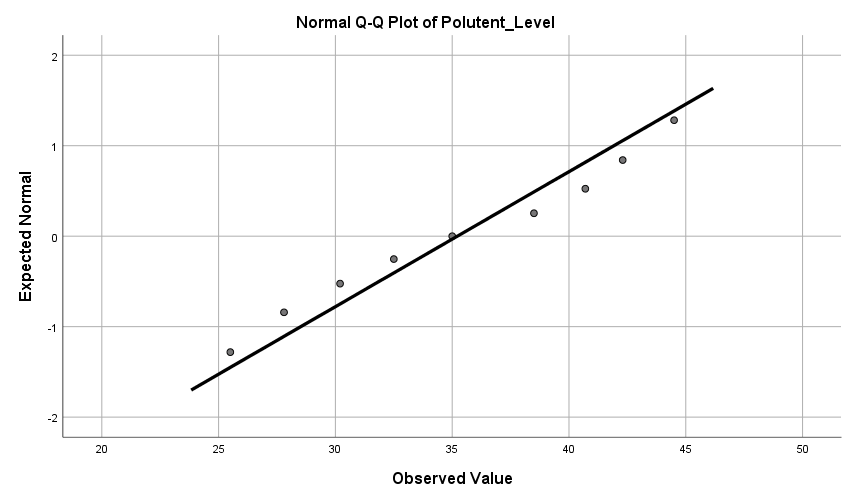
* **Purpose**: Evaluates whether the sample data is drawn from a normally distributed population.
* **Strength**: Particularly effective for small to medium-sized datasets.

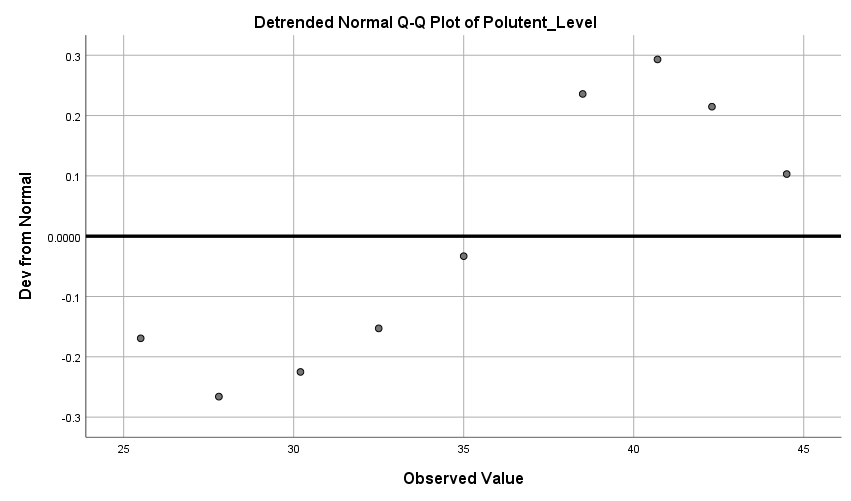
**Key Observations**

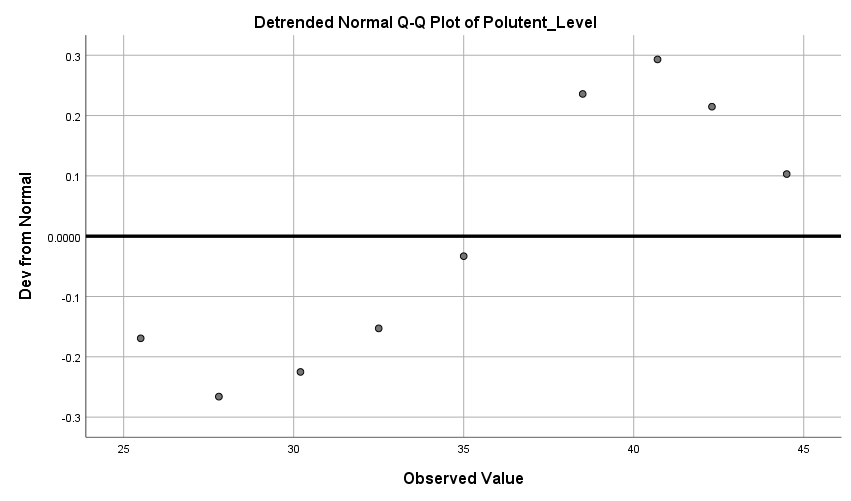
* **Consistent Results**: Both tests confirm that "Pollutant\_Level" and "Count\_of\_Affected\_Areas" are not normally distributed.
* **Recommendation**: Nonparametric statistical methods or data transformations should be considered in further analyses to accommodate this deviation from normality.

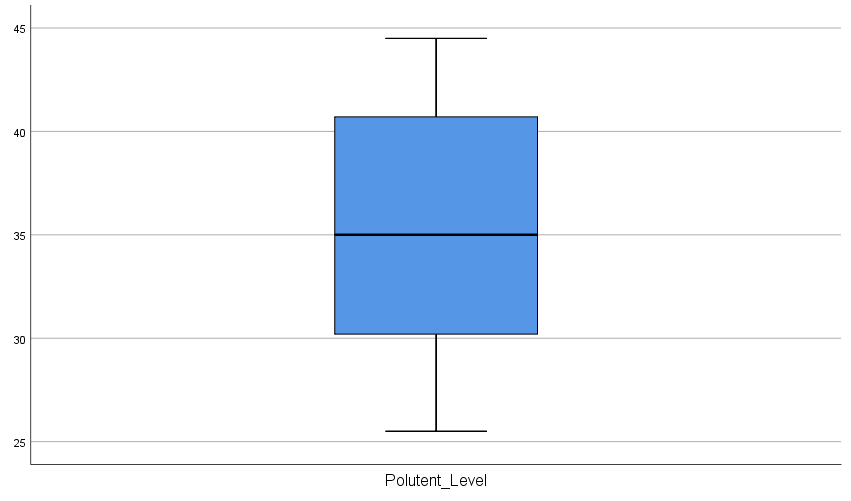
This analysis ensures appropriate statistical techniques are applied to achieve accurate results.

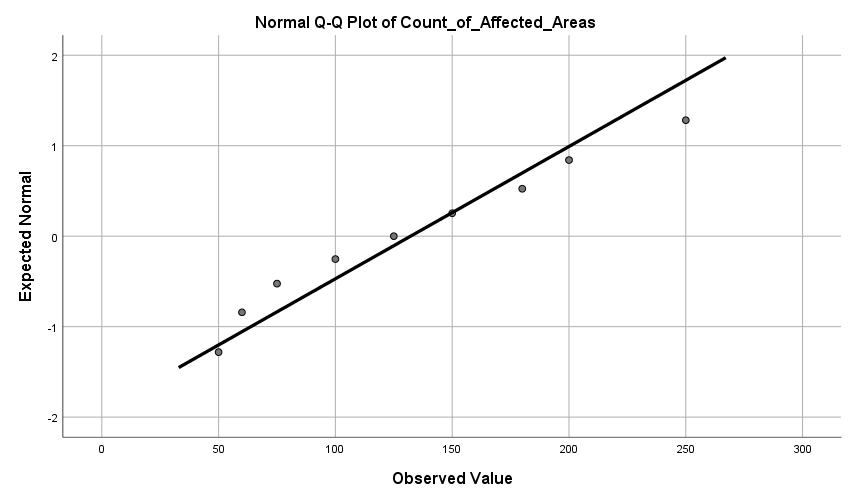
**Some Graph of Normality Test:**

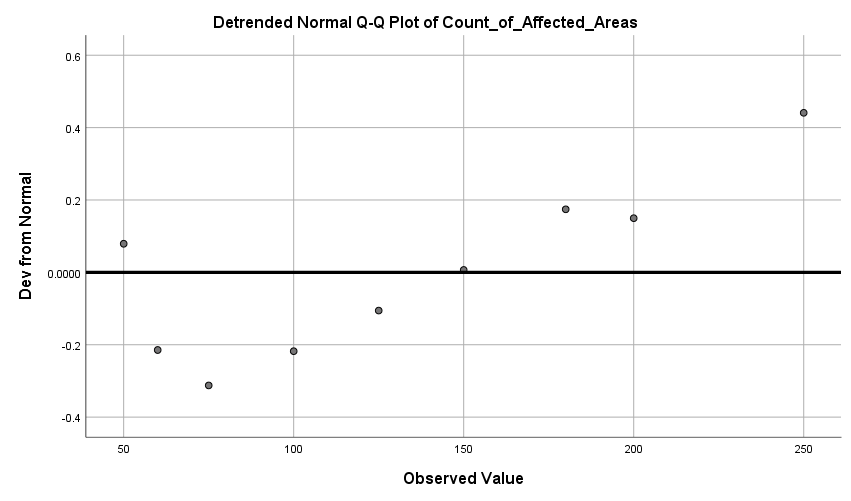
****

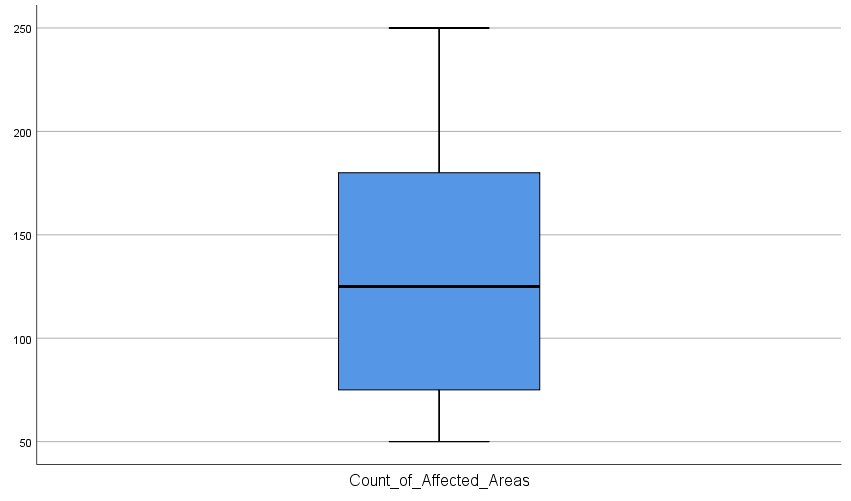
****

****

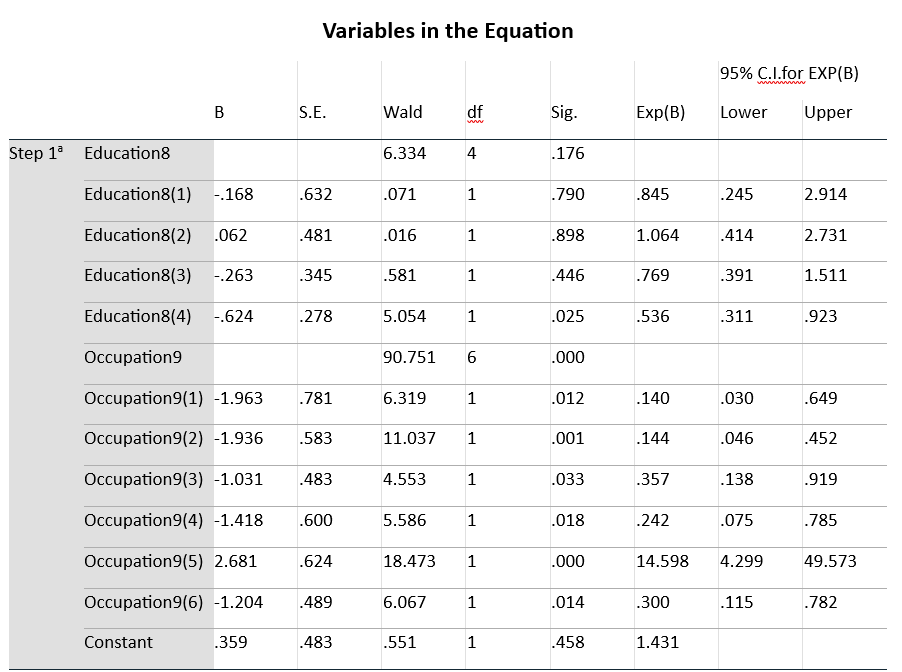
****

****

****

****

**Logistic Regression:**

****

**Figure 2:** Regression Analysis of Education and Occupation: Coefficients, Significance, and Confidence Intervals.

**Regression Analysis Summary**

This table presents the results of the regression analysis, detailing the relationship between predictor variables and the outcome variable. Key statistics include coefficients (B), standard errors (S.E.), Wald statistics, degrees of freedom (df), significance levels (Sig.), and 95% confidence intervals (C.I.).

**Key Predictor Variables**

1. **Education8**
   * Represents categorical education levels (e.g., levels 1 to 8).
   * Evaluates how education impacts the outcome variable.
2. **Occupation9**
   * Represents occupational categories (e.g., levels 1 to 9).
   * Assesses how different occupations influence the outcome variable.

**Key Metrics**

* **Coefficients (B):**  
  Indicate the direction and magnitude of the relationship:
  + **Positive Coefficients:** Direct relationship with the outcome variable.
  + **Negative Coefficients:** Inverse relationship with the outcome variable.
* **Standard Errors (S.E.):**  
  Measure the precision of the coefficients:
  + **Smaller Values:** Indicate more reliable estimates.
* **Wald Statistic:**  
  Tests the significance of predictors:
  + **Higher Values:** Suggest stronger contributions to the model.
* **Significance Levels (Sig.):**  
  Represent the probability of the observed relationship occurring by chance:
  + **Sig. < 0.05:** Indicates statistical significance.
* **Confidence Intervals (C.I.):**  
  Provide a range for the true value of the coefficient:
  + **C.I. excluding zero:** Indicates a significant effect.

**Analysis Overview**

* The regression analysis highlights the **influence of education and occupation** on the outcome variable.
* The metrics offer insights into the strength, direction, and significance of these relationships.
* This analysis aids in identifying meaningful patterns and supports informed decision-making.