**Module Code:** AQ049-3-M-MMDA

**Module Name:** Multivariate Methods for Data Analysis

# Individual Assignment

## Multiple Linear Regression & Factor Analysis (SPSS)


**Student Name:** Muhammad Yousouf Ali Budullah
**TP Number:** TP086704
**Programme:** MSc Data Science & Business Analytics
**Intake Code:** APDMF2501DSBA(BI)(PR)
**Module Lecturer:** Dr. Ang Chiew Way



**Date of Submission:** September 7, 2025

# Contents

# 1 Dataset Context & Metadata

## 1.1 Dataset Description

The dataset used in this study is the *Concrete Compressive Strength Dataset*, obtained from the UCI Machine Learning Repository (https://archive.ics.uci.edu/dataset/165/concrete+compressive+strength). It contains laboratory test results of concrete mixtures and their corresponding compressive strength values. The dataset consists of 1030 instances, each representing a mixture defined by its ingredient composition and curing age, with the compressive strength measured in megapascals (MPa).

The data is multivariate in nature and includes eight quantitative input variables representing the concrete ingredients and age, and one quantitative output variable representing the concrete compressive strength.

## 1.2 Metadata

The following table summarises the quantitative variables included in the dataset. At least one dependent variable and eight independent variables are included, satisfying the requirements for both Multiple Linear Regression (Q1) and Factor Analysis (Q2).

| **Variable Name (Label)** | **Type** | **SPSS Measure** | **Role (DV/IV)** |
|---|---|---|---|
| Cement (kg/m$^3$) | Numeric | Scale | Independent |
| Blast Furnace Slag (kg/m$^3$) | Numeric | Scale | Independent |
| Fly Ash (kg/m$^3$) | Numeric | Scale | Independent |
| Water (kg/m$^3$) | Numeric | Scale | Independent |
| Superplasticizer (kg/m$^3$) | Numeric | Scale | Independent |
| Coarse Aggregate (kg/m$^3$) | Numeric | Scale | Independent |
| Fine Aggregate (kg/m$^3$) | Numeric | Scale | Independent |
| Age (days) | Numeric | Scale | Independent |
| Concrete Compressive Strength (MPa) | Numeric | Scale | **Dependent Variable** |

Table 1: Metadata for the Concrete Compressive Strength Dataset (verified in SPSS).

## 1.3 Preprocessing

Minimal preprocessing was required as the dataset was already clean and free of missing values. All variables were confirmed as quantitative (Scale) in SPSS and kept for analysis.

# 2 Question 1: Multiple Linear Regression

## 2.1 Introduction (Q1a)

Concrete is one of the most commonly used construction materials in the world because of its its versatility, durability, and cost-effectiveness. Among its properties, compressive strength is considered the most crucial, as it directly determines the structural integrity and service life of concrete infrastructure. Ensuring the desired strength of concrete is important for safety, as subpar performance can result in structural failures with detrimental economic and human consequences. Traditional assessment methods rely on standardized compression tests, usually performed at 28 days, but these procedures are time-consuming, labor intensive, and sensitive to experimental errors (Sah & Hong, 2024).

To address these challenges, researchers have turned to predictive modeling approaches that estimate compressive strength based on input variables such as cement, aggregates, water, admixtures, and curing age. These models enable early strength estimation, which is helpful in making choices during construction and helps with mix optimisation before the complete curing period has passed. Predictive models present an opportunity to increase the efficiency and reliability of quality assurance processes by reducing the need for prolonged laboratory testing.

However, standard empirical models often struggle to accurately capture complex, nonlinear interactions between concrete components, limiting their effectiveness for prediction. As Nikoopayan Tak et al. (2025) states, such methods struggle to account for variability across mixes, illustrating the importance of strong computational tools that are able to accurately model these interactions. As a result, both statistical and machine learning approaches have been used in recent years, with multiple studies demonstrating that emerging methods surpass classical models in accuracy. Simpler regression-based models, on the other hand, continue to be useful because of their interpretability and ease of application in real-world engineering situations.

Building on this foundation, this study applies Multiple Linear Regression (MLR) to the UCI Concrete Compressive Strength dataset. By analysing the predictive influence of cement, slag, fly ash, water, superplasticizer, aggregates, and curing time, the analysis aims to identify the most significant determinants of compressive strength. In doing so, this study seeks to demonstrate how MLR can balance interpretability with predictive performance, providing practical information for data-driven improvements in concrete mix design and contributing to safer and more efficient construction practices

## 2.2 Problem Statement (Q1b)

A constant challenge in construction is the time and cost associated with measuring concrete compressive strength through standard laboratory tests, which typically require a 28-day curing period (Sah & Hong, 2024). This delay leaves projects vulnerable to safety and financial risks, as inadequate strength may only be identified after critical stages of construction.

The problem addressed in this study is the need for a reliable and interpretable predictive model that can estimate compressive strength from known proportions of the material and curing age. Multiple Linear Regression (MLR) is well suited for this task because as it quantifies the relationship between strength (dependent variable) and eight quantitative predictors (cement, slag, fly ash, water, superplasticizer, coarse aggregate, fine aggregate, and age).

Using the UCI Concrete Compressive Strength dataset with 1030 observations, this research applies MLR to identify the most significant predictors of strength and to provide a practical, data-driven tool that supports safer and more efficient decision-making in concrete mix design.

## 2.3 Research Objectives and Hypotheses (Q1c)

The aim of this study is to develop a Multiple Linear Regression (MLR) model to predict concrete compressive strength using quantitative variables from the UCI dataset. To achieve this aim, the following objectives and hypotheses are established.

- **Objective 1:** To evaluate the effect of material composition variables (cement, slag, fly ash, water, superplasticizer, coarse aggregate, and fine aggregate) on compressive strength.
  - $H_0$: Material composition variables have no significant effect on compressive strength.
  - $H_1$: At least one material composition variable has a significant effect on compressive strength.
- **Objective 2:** To examine the effect of curing age (days) on compressive strength.
  - $H_0$: Curing age has no significant effect on compressive strength.
  - $H_1$: Curing age has a significant effect on compressive strength.
- **Objective 3:** To test whether the combination of material composition and curing age collectively explains significant variance in compressive strength.
  - $H_0$: Material composition and curing age together do not significantly explain compressive strength.

– $H_1$: Material composition and curing age together significantly explain compressive strength.

## 2.4 Analysis Results and Interpretations (Q1d)

### 2.4.1 Dataset Evidence (Q1d-i)

Figure 1 shows the SPSS overview confirming that the dataset contains 1030 cases and 9 variables. All variables were defined as *Scale* (quantitative) in SPSS, as shown in Figure 2.

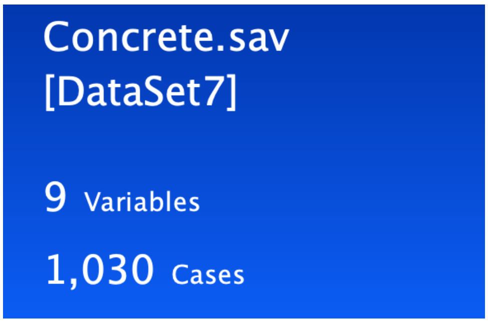


Figure 1: SPSS dataset overview confirming 1030 cases and 9 variables.

| | Name | Type | Width | Decimals | Label | Values | Missing | Columns | Align | Measure | Role |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | cement | Numeric | 8 | 2 | Cement (kg/m³) | None | None | 11 | Right | Scale | Input |
| 2 | slag | Numeric | 8 | 2 | Blast Furnace Slag (kg/m³) | None | None | 11 | Right | Scale | Input |
| 3 | flyash | Numeric | 8 | 2 | Fly Ash (kg/m³) | None | None | 11 | Right | Scale | Input |
| 4 | water | Numeric | 8 | 2 | Water (kg/m³) | None | None | 11 | Right | Scale | Input |
| 5 | superplast | Numeric | 8 | 3 | Superplasticizer (kg/m³) | None | None | 11 | Right | Scale | Input |
| 6 | coarseagg | Numeric | 8 | 2 | Coarse Aggregate (kg/m³) | None | None | 11 | Right | Scale | Input |
| 7 | fineagg | Numeric | 8 | 2 | Fine Aggregate (kg/m³) | None | None | 11 | Right | Scale | Input |
| 8 | age | Numeric | 11 | 0 | Curing Age (days) | None | None | 11 | Right | Scale | Input |
| 9 | strength | Numeric | 8 | 12 | Compressive Strength (MPa) | None | None | 11 | Right | Scale | Input |

Figure 2: SPSS Variable View showing all variables as quantitative (Scale).

### 2.4.2 Stepwise Multiple Linear Regression (Q1d-ii)

A stepwise multiple linear regression was conducted in SPSS with *Compressive Strength (MPa)* as the dependent variable. The independent variables included eight quantitative predictors: cement (kg/m$^3$), blast furnace slag (kg/m$^3$), fly ash (kg/m$^3$), water (kg/m$^3$), superplasticizer (kg/m$^3$), coarse aggregate (kg/m$^3$), fine aggregate (kg/m$^3$), and curing age (days). These variables represent the mixture proportions and curing conditions of concrete. SPSS applied the stepwise method to retain only statistically significant predictors that improved the overall model fit, with the detailed results presented in the following subsections.

### 2.4.3 Assumption Testing (Q1d-iii)

To ensure the validity of the regression results, the four key assumptions of multiple linear regression were examined: normality, linearity, homoscedasticity, and multicollinearity.

## Normality of residuals

As the sample size exceeded 50 ($N$ = 1030), the Kolmogorov–Smirnov test was considered more appropriate than the Shapiro–Wilk test for assessing normality. The result ( p = .025) indicated a statistically significant deviation from perfect normality. However, the Normal P–P plot of the standardized residuals (Figure 3) showed points closely followed the diagonal line, suggesting that the residuals were approximately normally distributed. Therefore, the normality assumption was considered satisfied.

**Tests of Normality**

| | Kolmogorov-Smirnov[a] | | | Shapiro-Wilk | | |
|---|---|---|---|---|---|---|
| | Statistic | df | Sig. | Statistic | df | Sig. |
| Unstandardized Residual | .031 | 1030 | .025 | .996 | 1030 | .007 |

a. Lilliefors Significance Correction

Table 2: Kolmogorov–Smirnov and Shapiro–Wilk tests of normality for regression residuals.

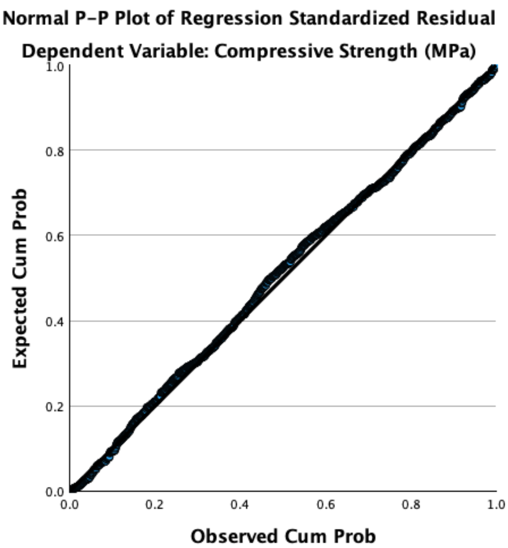


Figure 3: Normal P–P plot of regression standardized residuals (Model 6).

## Linearity and Homoscedasticity

The scatterplot of standardized residuals against predicted values (Figure 4) displayed a random pattern without systematic curvature or U-shapes, confirming linearity. The spread of residuals was relatively constant across predicted values, with no funnel-shaped pattern present, indicating that the assumption of homoscedasticity (equal variance of errors) was also satisfied.

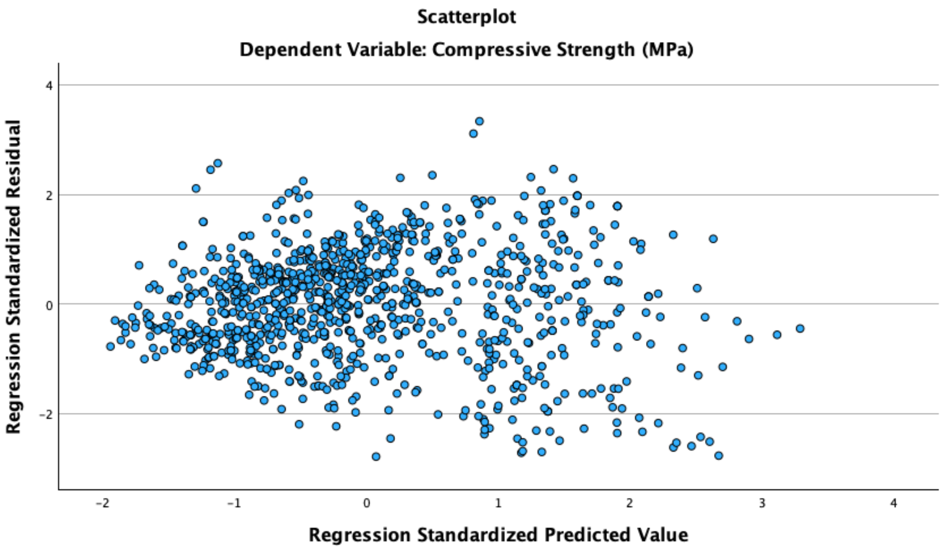


Figure 4: Scatterplot of standardized residuals versus predicted values (Model 6).

## Multicollinearity

The coefficients output for the final regression model (Model 6) reported Variance Inflation Factor (VIF) values all below 2.5 and tolerance values above 0.1, as shown in Table 3. These results confirm that multicollinearity was not a concern in the final regression model.

**Coefficients[a]**

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. | 95.0% Confidence Interval for B | | Collinearity Statistics | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | B | Std. Error | Beta | | | Lower Bound | Upper Bound | Tolerance | VIF |
| 1 | (Constant) | 13.443 | 1.297 | | 10.365 | <.001 | 10.898 | 15.988 | | |
| | Cement (kg/m³) | .080 | .004 | .498 | 18.405 | <.001 | .071 | .088 | 1.000 | 1.000 |
| 2 | (Constant) | 9.190 | 1.250 | | 7.350 | <.001 | 6.737 | 11.644 | | |
| | Cement (kg/m³) | .075 | .004 | .468 | 18.534 | <.001 | .067 | .083 | .991 | 1.009 |
| | Superplasticizer (kg/m³) | .902 | .071 | .323 | 12.782 | <.001 | .764 | 1.041 | .991 | 1.009 |
| 3 | (Constant) | 5.098 | 1.147 | | 4.446 | <.001 | 2.848 | 7.348 | | |
| | Cement (kg/m³) | .069 | .004 | .431 | 18.976 | <.001 | .062 | .076 | .981 | 1.019 |
| | Superplasticizer (kg/m³) | 1.112 | .064 | .397 | 17.246 | <.001 | .985 | 1.238 | .951 | 1.052 |
| | Curing Age (days) | .098 | .006 | .370 | 16.077 | <.001 | .086 | .110 | .953 | 1.050 |
| 4 | (Constant) | -2.001 | 1.208 | | -1.657 | .098 | -4.371 | .369 | | |
| | Cement (kg/m³) | .081 | .004 | .508 | 23.076 | <.001 | .074 | .088 | .905 | 1.105 |
| | Superplasticizer (kg/m³) | 1.060 | .060 | .379 | 17.610 | <.001 | .941 | 1.178 | .946 | 1.057 |
| | Curing Age (days) | .098 | .006 | .372 | 17.368 | <.001 | .087 | .110 | .953 | 1.050 |
| | Blast Furnace Slag (kg/m³) | .053 | .004 | .275 | 12.580 | <.001 | .045 | .061 | .919 | 1.088 |
| 5 | (Constant) | 35.492 | 4.304 | | 8.246 | <.001 | 27.046 | 43.938 | | |
| | Cement (kg/m³) | .081 | .003 | .509 | 24.007 | <.001 | .075 | .088 | .905 | 1.105 |
| | Superplasticizer (kg/m³) | .614 | .076 | .219 | 8.073 | <.001 | .465 | .763 | .549 | 1.821 |
| | Curing Age (days) | .109 | .006 | .413 | 19.554 | <.001 | .098 | .120 | .910 | 1.100 |
| | Blast Furnace Slag (kg/m³) | .060 | .004 | .311 | 14.520 | <.001 | .052 | .068 | .888 | 1.127 |
| | Water (kg/m³) | -.197 | .022 | -.252 | -9.047 | <.001 | -.240 | -.154 | .524 | 1.908 |
| 6 | (Constant) | 29.030 | 4.212 | | 6.891 | <.001 | 20.764 | 37.296 | | |
| | Cement (kg/m³) | .105 | .004 | .660 | 24.821 | <.001 | .097 | .114 | .534 | 1.871 |
| | Superplasticizer (kg/m³) | .239 | .085 | .085 | 2.826 | .005 | .073 | .405 | .412 | 2.424 |
| | Curing Age (days) | .113 | .005 | .429 | 20.987 | <.001 | .103 | .124 | .902 | 1.108 |
| | Blast Furnace Slag (kg/m³) | .086 | .005 | .447 | 17.386 | <.001 | .077 | .096 | .572 | 1.749 |
| | Water (kg/m³) | -.218 | .021 | -.279 | -10.332 | <.001 | -.260 | -.177 | .517 | 1.933 |
| | Fly Ash (kg/m³) | .069 | .008 | .263 | 8.881 | <.001 | .054 | .084 | .430 | 2.328 |

a. Dependent Variable: Compressive Strength (MPa)

Table 3: Collinearity statistics (Tolerance and VIF) for predictors in Model 6.

### 2.4.4 Model Summary Interpretation (Q1d-iv)

Table 4 presents the regression progress across six steps. The final model (Model 6) achieved $R = 0.784$ and an adjusted $R^2 = 0.612$, indicating that the predictors collectively explain 61.2% of the variance in compressive strength. The standard error of the estimate decreased from 14.50 in the initial model to 10.41 in the final model, confirming the improved predictive accuracy as relevant predictors were retained.

These findings address Objective 3, showing that material composition and curing age jointly account for a substantial proportion of variability in compressive strength. The null hypothesis ($H_0$) that these predictors do not explain compressive strength is therefore rejected.

**Model Summary[g]**

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
|---|---|---|---|---|
| 1 | .498[a] | .248 | .247 | 14.495430857581399 |
| 2 | .593[b] | .351 | .350 | 13.470522571358584 |
| 3 | .694[c] | .482 | .480 | 12.045095584476996 |
| 4 | .742[d] | .551 | .549 | 11.216192566254911 |
| 5 | .764[e] | .584 | .582 | 10.798373985681206 |
| 6 | .784[f] | .614 | .612 | 10.409784091934055 |

a. Predictors: (Constant), Cement (kg/m³)
b. Predictors: (Constant), Cement (kg/m³), Superplasticizer (kg/m³)
c. Predictors: (Constant), Cement (kg/m³), Superplasticizer (kg/m³), Curing Age (days)
d. Predictors: (Constant), Cement (kg/m³), Superplasticizer (kg/m³), Curing Age (days), Blast Furnace Slag (kg/m³)
e. Predictors: (Constant), Cement (kg/m³), Superplasticizer (kg/m³), Curing Age (days), Blast Furnace Slag (kg/m³), Water (kg/m³)
f. Predictors: (Constant), Cement (kg/m³), Superplasticizer (kg/m³), Curing Age (days), Blast Furnace Slag (kg/m³), Water (kg/m³), Fly Ash (kg/m³)
g. Dependent Variable: Compressive Strength (MPa)

Table 4: Model summary of stepwise regression predicting compressive strength.

### 2.4.5 Model Adequacy and Significance (Q1d-v)

**Hypotheses:**

- $H_0$: The model does not fit the data.
- $H_1$: The model fits the data.

**ANOVA**[a]

| Model | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| 1 | Regression | 71172.222 | 1 | 71172.222 | 338.726 | <.001[b] |
| | Residual | 216000.806 | 1028 | 210.118 | | |
| | Total | 287173.028 | 1029 | | | |
| 2 | Regression | 100818.766 | 2 | 50409.383 | 277.807 | <.001[c] |
| | Residual | 186354.263 | 1027 | 181.455 | | |
| | Total | 287173.028 | 1029 | | | |
| 3 | Regression | 138316.508 | 3 | 46105.503 | 317.784 | <.001[d] |
| | Residual | 148856.520 | 1026 | 145.084 | | |
| | Total | 287173.028 | 1029 | | | |
| 4 | Regression | 158224.978 | 4 | 39556.245 | 314.430 | <.001[e] |
| | Residual | 128948.050 | 1025 | 125.803 | | |
| | Total | 287173.028 | 1029 | | | |
| 5 | Regression | 167769.631 | 5 | 33553.926 | 287.757 | <.001[f] |
| | Residual | 119403.398 | 1024 | 116.605 | | |
| | Total | 287173.028 | 1029 | | | |
| 6 | Regression | 176317.061 | 6 | 29386.177 | 271.181 | <.001[g] |
| | Residual | 110855.968 | 1023 | 108.364 | | |
| | Total | 287173.028 | 1029 | | | |

a. Dependent Variable: Compressive Strength (MPa)
b. Predictors: (Constant), Cement (kg/m³)
c. Predictors: (Constant), Cement (kg/m³), Superplasticizer (kg/m³)
d. Predictors: (Constant), Cement (kg/m³), Superplasticizer (kg/m³), Curing Age (days)
e. Predictors: (Constant), Cement (kg/m³), Superplasticizer (kg/m³), Curing Age (days), Blast Furnace Slag (kg/m³)
f. Predictors: (Constant), Cement (kg/m³), Superplasticizer (kg/m³), Curing Age (days), Blast Furnace Slag (kg/m³), Water (kg/m³)
g. Predictors: (Constant), Cement (kg/m³), Superplasticizer (kg/m³), Curing Age (days), Blast Furnace Slag (kg/m³), Water (kg/m³), Fly Ash (kg/m³)

Table 5: ANOVA results for stepwise regression predicting compressive strength (Model 6).

The ANOVA results in Table 5 confirm the statistical adequacy of the regression model. For the final stepwise model (Model 6), the regression sum of squares was 176,317.061 compared to a residual sum of 110,855.968, yielding an F-statistic of 271.181 with $df = (6, 1023)$ and a significance level of $p < 0.001$.

This leads to the rejection of $H_0$ as there is sufficient evidence to conclude that the independent variables (cement, burst furnace slag, fly ash, water, superplasticizer, and curing age) collectively provide a significant prediction of compressive strength. Therefore, the regression model is a good fit for the data and directly addresses Objective 3.

### 2.4.6 Coefficient Interpretation (Q1d-vi

**Coefficients[a]**

| Model | | Unstandardized Coefficients B | Std. Error | Standardized Coefficients Beta | t | Sig. | 95.0% Confidence Interval for B Lower Bound | Upper Bound | Collinearity Statistics Tolerance | VIF |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | (Constant) | 13.443 | 1.297 | | 10.365 | <.001 | 10.898 | 15.988 | | |
| | Cement (kg/m³) | .080 | .004 | .498 | 18.405 | <.001 | .071 | .088 | 1.000 | 1.000 |
| 2 | (Constant) | 9.190 | 1.250 | | 7.350 | <.001 | 6.737 | 11.644 | | |
| | Cement (kg/m³) | .075 | .004 | .468 | 18.534 | <.001 | .067 | .083 | .991 | 1.009 |
| | Superplasticizer (kg/m³) | .902 | .071 | .323 | 12.782 | <.001 | .764 | 1.041 | .991 | 1.009 |
| 3 | (Constant) | 5.098 | 1.147 | | 4.446 | <.001 | 2.848 | 7.348 | | |
| | Cement (kg/m³) | .069 | .004 | .431 | 18.976 | <.001 | .062 | .076 | .981 | 1.019 |
| | Superplasticizer (kg/m³) | 1.112 | .064 | .397 | 17.246 | <.001 | .985 | 1.238 | .951 | 1.052 |
| | Curing Age (days) | .098 | .006 | .370 | 16.077 | <.001 | .086 | .110 | .953 | 1.050 |
| 4 | (Constant) | -2.001 | 1.208 | | -1.657 | .098 | -4.371 | .369 | | |
| | Cement (kg/m³) | .081 | .004 | .508 | 23.076 | <.001 | .074 | .088 | .905 | 1.105 |
| | Superplasticizer (kg/m³) | 1.060 | .060 | .379 | 17.610 | <.001 | .941 | 1.178 | .946 | 1.057 |
| | Curing Age (days) | .098 | .006 | .372 | 17.368 | <.001 | .087 | .110 | .953 | 1.050 |
| | Blast Furnace Slag (kg/m³) | .053 | .004 | .275 | 12.580 | <.001 | .045 | .061 | .919 | 1.088 |
| 5 | (Constant) | 35.492 | 4.304 | | 8.246 | <.001 | 27.046 | 43.938 | | |
| | Cement (kg/m³) | .081 | .003 | .509 | 24.007 | <.001 | .075 | .088 | .905 | 1.105 |
| | Superplasticizer (kg/m³) | .614 | .076 | .219 | 8.073 | <.001 | .465 | .763 | .549 | 1.821 |
| | Curing Age (days) | .109 | .006 | .413 | 19.554 | <.001 | .098 | .120 | .910 | 1.100 |
| | Blast Furnace Slag (kg/m³) | .060 | .004 | .311 | 14.520 | <.001 | .052 | .068 | .888 | 1.127 |
| | Water (kg/m³) | -.197 | .022 | -.252 | -9.047 | <.001 | -.240 | -.154 | .524 | 1.908 |
| 6 | (Constant) | 29.030 | 4.212 | | 6.891 | <.001 | 20.764 | 37.296 | | |
| | Cement (kg/m³) | .105 | .004 | .660 | 24.821 | <.001 | .097 | .114 | .534 | 1.871 |
| | Superplasticizer (kg/m³) | .239 | .085 | .085 | 2.826 | .005 | .073 | .405 | .412 | 2.424 |
| | Curing Age (days) | .113 | .005 | .429 | 20.987 | <.001 | .103 | .124 | .902 | 1.108 |
| | Blast Furnace Slag (kg/m³) | .086 | .005 | .447 | 17.386 | <.001 | .077 | .096 | .572 | 1.749 |
| | Water (kg/m³) | -.218 | .021 | -.279 | -10.332 | <.001 | -.260 | -.177 | .517 | 1.933 |
| | Fly Ash (kg/m³) | .069 | .008 | .263 | 8.881 | <.001 | .054 | .084 | .430 | 2.328 |

a. Dependent Variable: Compressive Strength (MPa)

Table 6: Coefficients table for stepwise regression predicting compressive strength (Model 6).

**Objective 1: Effect of material composition**

The final model (Model 6) in Figure 6 shows that cement ($\beta = 0.660$, $p < .001$) was the strongest predictor of compressive strength. Blast furnace slag ($\beta = 0.447$, $p < .001$) and fly ash ($\beta = 0.263$, $p < .001$) also improved strength, reflecting their pozzolanic contribution. Superplasticizer had a smaller positive effect ($p = .005$), while water reduced strength ($\beta = -0.279$, $p < .001$), in line with the weakening effect of excess water on concrete. Aggregates were excluded, indicating their limited influence once other variables were controlled.

These results are consistent with the existing literature. Sah and Hong (2024). found that cement content had the strongest positive association with compressive strength, while water content showed a clear negative effect, confirming the dominant role of the water–cement ratio. Similarly, Nikoopayan Tak et al. (2025) reported that additional cementitious materials such as blast furnace slag and fly ash improve strength through pozzolanic reactions, with superplasticizer providing additional but smaller gains.

**Objective 2: Effect of curing age**

Curing age ($\beta = 0.429$, $p < .001$) was also significant, confirming that strength increases with time through continued hydration. This supports the hypothesis and agrees with Nikoopayan Tak et al. (2025), who observed consistent strength gains with longer curing periods and noted that their results were in line with prior investigations reporting similar trends.

## 2.5 Conclusions and Recommendations (Q1e)

### Conclusions

This study demonstrated that multiple linear regression can effectively predict concrete compressive strength from mixture proportions and curing age. The final stepwise model explained 61.2% of the variance, identifying cement content, curing age, slag, and fly ash as significant positive predictors, while water had a negative effect. These results confirmed the hypotheses for all the objectives and were consistent with previous research emphasising the central role of cementitious materials and curing conditions in determining compressive strength .

### Recommendations

The findings highlight practical approaches for companies in the industry. Strength can be increased by optimising cement and water ratios, extending curing periods, and incorporating supplementary cementitious materials such as slag and fly ash. Water content should be carefully controlled, while superplasticizers can be used to improve workability without compromising performance.

### Limitations & Future Research

It is important to note this study relied on a secondary dataset generated under laboratory conditions, which may not capture field variability such as temperature, humidity, or long-term durability. Furthermore, stepwise regression may exclude variables that could be relevant in other contexts, and the linear approach may not fully account for complex material interactions. Future research could expand on this work by using larger and more diverse datasets, including field measurements. Exploring non-linear and machine learning methodologies, as well as integrating durability and sustainability indicators, could provide a more comprehensive understanding of concrete performance.

# 3 Question 2: Factor Analysis

## 3.1 Purpose of Factor Analysis (Q2a)

Factor analysis is a multivariate technique used to identify underlying structures among a set of observed variables. It is particularly appropriate for this study because the dataset contains several correlated predictors (cement, slag, fly ash, water, superplasticizer, coarse aggregate, fine aggregate, and curing age), which may reflect latent dimensions such as "cementitious materials," "aggregates," or "curing conditions."

The main purpose of factor analysis is to reduce dimensionality while retaining most of the variance present in the original dataset. By grouping correlated variables into common factors, factor analysis simplifies complex datasets, reduces redundancy, and allows a clearer interpretation of how different inputs relate to one another. In this context, it helps to uncover the main components that infleunce compressive strength, ensuring that subsequent analyses focus on interpretable components rather than individual overlapping predictors.

## 3.2 Handling of Non-Metric Variables (Q2b)

All independent variables used in this study are metric and measured on a scale level, which makes them appropriate for factor analysis. Consequently, no non-metric variables were included in the procedure.

Factor analysis assumes continuous, normally distributed data and linear relationships among variables. Non-metric variables, such as categorical or nominal data, are generally excluded because they violate these assumptions and can distort the factor structure. If such variables were present, they would either be removed or transformed into dummy or ordinal-coded variables to allow for an appropriate analysis.

## 3.3 Execution of Factor Analysis (Q2c)

To evaluate the structure of the independent variables, factor analysis was conducted using Principal Component Analysis (PCA) with Varimax rotation. The following SPSS outputs provide evidence of sampling adequacy, factor extraction, and loadings.

## Communalities

| | Initial | Extraction |
|---|---|---|
| Cement (kg/m³) | 1.000 | .932 |
| Blast Furnace Slag (kg/m³) | 1.000 | .911 |
| Fly Ash (kg/m³) | 1.000 | .659 |
| Water (kg/m³) | 1.000 | .836 |
| Superplasticizer (kg/m³) | 1.000 | .772 |
| Coarse Aggregate (kg/m³) | 1.000 | .908 |
| Fine Aggregate (kg/m³) | 1.000 | .520 |
| Curing Age (days) | 1.000 | .512 |

Extraction Method: Principal Component Analysis.

Table 8: Communalities table showing proportion of variance explained for each variable.

**Total Variance Explained**

| Component | Initial Eigenvalues | | | Extraction Sums of Squared Loadings | | | Rotation Sums of Squared Loadings | | |
|---|---|---|---|---|---|---|---|---|---|
| | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % |
| 1 | 2.280 | 28.499 | 28.499 | 2.280 | 28.499 | 28.499 | 1.951 | 24.385 | 24.385 |
| 2 | 1.416 | 17.703 | 46.201 | 1.416 | 17.703 | 46.201 | 1.429 | 17.860 | 42.246 |
| 3 | 1.340 | 16.753 | 62.954 | 1.340 | 16.753 | 62.954 | 1.415 | 17.692 | 59.938 |
| 4 | 1.014 | 12.677 | 75.631 | 1.014 | 12.677 | 75.631 | 1.255 | 15.694 | 75.631 |
| 5 | .952 | 11.895 | 87.526 | | | | | | |
| 6 | .790 | 9.877 | 97.403 | | | | | | |
| 7 | .178 | 2.222 | 99.624 | | | | | | |
| 8 | .030 | .376 | 100.000 | | | | | | |

Extraction Method: Principal Component Analysis.

Table 9: Total variance explained by extracted components.

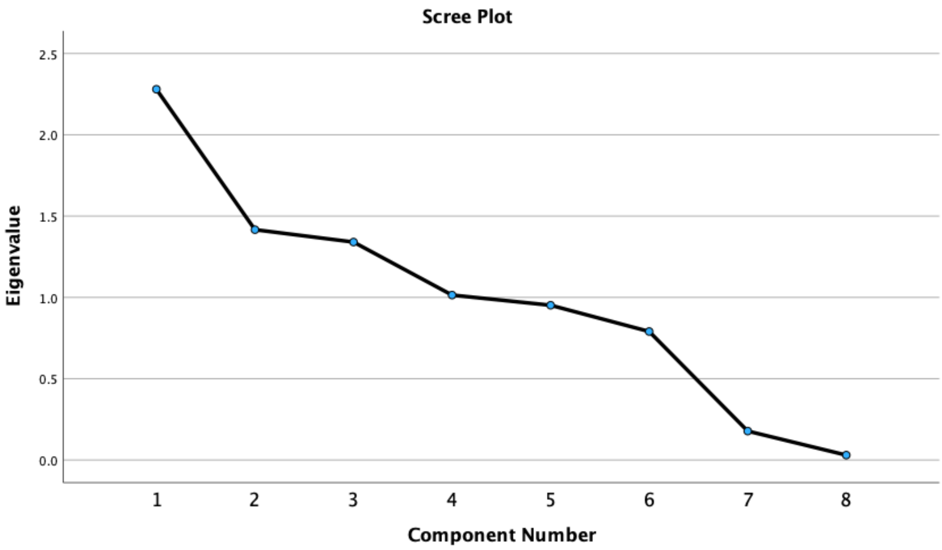


Figure 5: Scree plot of eigenvalues indicating factor retention.

**Component Matrix[a]**

| | Component | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| Cement (kg/m³) | -.148 | .133 | .943 | -.055 |
| Blast Furnace Slag (kg/m³) | -.268 | -.816 | -.201 | -.365 |
| Fly Ash (kg/m³) | .596 | .171 | -.472 | .228 |
| Water (kg/m³) | -.826 | -.063 | -.247 | .298 |
| Superplasticizer (kg/m³) | .764 | -.338 | .271 | -.038 |
| Coarse Aggregate (kg/m³) | -.057 | .750 | -.200 | -.550 |
| Fine Aggregate (kg/m³) | .607 | .023 | -.006 | .388 |
| Curing Age (days) | -.440 | .150 | .117 | .532 |

Extraction Method: Principal Component Analysis.

a. 4 components extracted.

Table 10: Unrotated component matrix.

**Rotated Component Matrix[a]**

| | Component | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| Cement (kg/m³) | .084 | .940 | .196 | .047 |
| Blast Furnace Slag (kg/m³) | .033 | -.076 | -.922 | .234 |
| Fly Ash (kg/m³) | .269 | -.662 | .384 | .032 |
| Water (kg/m³) | -.878 | -.016 | -.247 | .056 |
| Superplasticizer (kg/m³) | .781 | -.003 | .094 | .391 |
| Coarse Aggregate (kg/m³) | .028 | -.041 | .152 | -.939 |
| Fine Aggregate (kg/m³) | .334 | -.265 | .470 | .342 |
| Curing Age (days) | -.613 | .168 | .257 | .205 |

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 7 iterations.

Table 11: Rotated component matrix (Varimax rotation).

**Component Transformation Matrix**

| Component | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | .848 | -.322 | .389 | .163 |
| 2 | -.170 | .045 | .700 | -.693 |
| 3 | .220 | .930 | .203 | .212 |
| 4 | -.452 | -.169 | .564 | .670 |

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

Table 12: Component transformation matrix.

## 3.4 Grouping of Variables (Q2d)

**Rotated Component Matrix[a]**

| | Component | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| Cement (kg/m³) | .084 | .940 | .196 | .047 |
| Blast Furnace Slag (kg/m³) | .033 | -.076 | -.922 | .234 |
| Fly Ash (kg/m³) | .269 | -.662 | .384 | .032 |
| Water (kg/m³) | -.878 | -.016 | -.247 | .056 |
| Superplasticizer (kg/m³) | .781 | -.003 | .094 | .391 |
| Coarse Aggregate (kg/m³) | .028 | -.041 | .152 | -.939 |
| Fine Aggregate (kg/m³) | .334 | -.265 | .470 | .342 |
| Curing Age (days) | -.613 | .168 | .257 | .205 |

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 7 iterations.

Table 13: Rotated component matrix (Grouping).

■ **Factor 1: Water–Admixture–Curing**

Water (-0.878), Superplasticizer (0.781), and Curing Age (-0.613) clustered together, representing the balance between mix fluidity, chemical admixtures, and time-dependent hydration.

■ **Factor 2: Primary Binder**

Cement (0.940) and Fly Ash (-0.662) loaded strongly, reflecting the role of cement and its partial replacement material in strength development.

■ **Factor 3: Supplementary Binder and Fines**

Blast Furnace Slag (-0.922) and Fine Aggregate (0.470) grouped under this factor, highlighting the contribution of secondary binders and finer particles to concrete properties.

■ **Factor 4: Coarse Aggregate**

Coarse Aggregate (-0.939) loaded almost exclusively on this factor, capturing the structural skeleton of the mix.

## 3.5 Interpretation of Outputs (Q2e)

### 3.5.1 Interpretation of Communalities (Q2i)

As shown in Table 8, Cement (0.932), Blast Furnace Slag (0.911), and Coarse Aggregate (0.908) had very high extraction values, meaning over 90% of their variance was explained by the extracted factors. Water (0.836) and Superplasticizer (0.772) were also well represented.

Fly Ash (0.659), Fine Aggregate (0.520), and Curing Age (0.512) had lower values, but still above the recommended 0.50 threshold. This indicates that just over half of their variance was explained by the model, which is considered adequate for retention.

According to the factor analysis rule of thumb, communalities below 0.50 suggest removing a variable. However, since all variables here exceeded 0.50, none required exclusion from the analysis.

### 3.5.2 Interpretation of Eigenvalues and Scree Plot (Q2ii)

The eigenvalue results (Table 9) show that the first four components had values greater than 1 (2.280, 1.416, 1.340, and 1.014), together accounting for 75.6% of the total variance. According to the Kaiser criterion, these four components should be retained, while the remaining components with eigenvalues below 1 contribute minimal explanatory power.

The scree plot (Figure 5) displays a clear decline in eigenvalues across the first four components, after which the values drop below 1 and level off. This pattern confirms that retaining four factors provides a balanced representation of the dataset. Thus, both the Kaiser rule and scree test support a four-factor solution, consistent with the rotated component matrix (Table 13).

## 3.6 Improving Factorability & Reliability (Q2f-i)

Reliability could be improved in two ways. First, variables with relatively low communalities, such as Fine Aggregate (.520) and Curing Age (.512), may weaken the stability of the factor solution and should be reconsidered or measured with greater precision in future analyses.

Second, although the large sample size (N = 1030) already exceeds the guidelines, which improves the stability of the factor solution. Research shows that variables with high communalities need fewer cases, while weaker variables require larger samples to achieve accurate results (Lorenzo-Seva & Ferrando, 2024). This confirms that the dataset is strong, though addressing weaker variables could further improve its reliability.

The robustness of factor solutions depends strongly on both sample size and variable quality. Larger samples help stabilize communalities and factor loadings, making the solution less sensitive to sampling error. At the same time, including variables with weak or inconsistent loadings can introduce noise, reducing clarity and reliability. A balance of sufficient cases and carefully selected variables therefore provides the most stable and interpretable factor structure.

## 3.7 Cross-Loadings (Q2f-ii, iii)

### 3.7.1 Understanding and Addressing Cross-Loadings

A cross-loading occurs when a variable loads substantially on more than one factor, making it unclear which construct it belongs to. Thresholds above 0.40 on multiple components are commonly treated as problematic.

In the unrotated component matrix (Table 10), Fly Ash loaded 0.596 on Component 1 and –0.472 on Component 3, Coarse Aggregate loaded 0.750 on Component 2 and -0.550 on Component 4 and Curing age loaded -0.440 on Component 1 and 0.532 on Component 4 exceeding the 0.40 threshold, indicating cross-loadings. After Varimax rotation (Table 11), the issue was reduced but not fully eliminated. Fly Ash (–0.662 on Component 2; 0.384 on Component 3) Fine Aggregate (0.470 on Component 3; 0.334 on Component 1) and Superplasticizer (0.781 on Component 1; 0.391 on Component 4) showed moderate loadings on more than one factor. Although one loading in each case fell just below the 0.40 threshold, it is still large enough to suggest an influence from two factors.

Such cross-loadings can distort interpretation by blurring the meaning of the components. For example, Fly Ash may appear to represent both the binder dimension and the fines dimension, however this limits the ability to treat each factor as a distinct component.

### 3.7.2 Reducing Cross-Loadings

Cross-loadings reduces the clarity of factor analysis because a variable seems to belong to more than one factor. Two main strategies can be applied. First, variables that show consistent overlap, such as Fly Ash and Fine Aggregate in this dataset, can be reconsidered or removed, since excluding vague items improves the stability of the solution. Second, the choice of rotation method can reduce cross-loadings.

While Varimax is the most common rotation, it is sensitive to sampling error because it minimises individual cross-loadings. Beauducel and Hilger (2023) demonstrated that oblique mean target (OMT) rotation, which minimises the average cross-loadings across factors, produces more robust and interpretable results. This suggests that in future analyses, cross-loadings could be reduced either by refining weak variables or by using more flexible rotation techniques such as OMT, which may provide clearer factor solutions than Varimax.

# References

Beauducel, A., & Hilger, N. (2023). Robust oblique Target-rotation for small samples. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1285212

Lorenzo-Seva, U., & Ferrando, P. J. (2024). Determining Sample Size Requirements in EFA Solutions: A Simple Empirical Proposal. *Multivariate Behavioral Research*, *59*(5), 899–912. https://doi.org/10.1080/00273171.2024.2342324

Nikoopayan Tak, M. S., Feng, Y., & Mahgoub, M. (2025). Advanced Machine Learning Techniques for Predicting Concrete Compressive Strength. *Infrastructures*, *10*(2). https://doi.org/10.3390/infrastructures10020026

Sah, A. K., & Hong, Y. M. (2024). Performance Comparison of Machine Learning Models for Concrete Compressive Strength Prediction. *Materials*, *17*(9). https://doi.org/10.3390/ma17092075

# Appendices

## Appendix A: SPSS Outputs for Q1 (Regression)

# Appendix A: Multiple Linear Regression Outputs

This appendix presents the complete SPSS outputs for the multiple linear regression analysis. All figures are labelled for reference in Section 1(d).

**Descriptive Statistics**

| | Mean | Std. Deviation | N |
|---|---|---|---|
| Compressive Strength (MPa) | 35.81783582611362 | 16.705679174867940 | 1030 |
| Cement (kg/m³) | 281.1656 | 104.50714 | 1030 |
| Blast Furnace Slag (kg/m³) | 73.8955 | 86.27910 | 1030 |
| Fly Ash (kg/m³) | 54.1871 | 63.99647 | 1030 |
| Water (kg/m³) | 181.5664 | 21.35557 | 1030 |
| Superplasticizer (kg/m³) | 6.20311 | 5.973492 | 1030 |
| Coarse Aggregate (kg/m³) | 972.9186 | 77.75382 | 1030 |
| Fine Aggregate (kg/m³) | 773.5789 | 80.17543 | 1030 |
| Curing Age (days) | 45.66 | 63.170 | 1030 |

Table 14: Descriptive statistics of dependent and independent variables.

**Correlations**

| | | Compressive Strength (MPa) | Cement (kg/m³) | Blast Furnace Slag (kg/m³) | Fly Ash (kg/m³) | Water (kg/m³) | Superplasticizer (kg/m³) | Coarse Aggregate (kg/m³) | Fine Aggregate (kg/m³) | Curing Age (days) |
|---|---|---|---|---|---|---|---|---|---|---|
| Pearson Correlation | Compressive Strength (MPa) | 1.000 | .498 | .135 | -.106 | -.290 | .366 | -.165 | -.167 | .329 |
| | Cement (kg/m³) | .498 | 1.000 | -.275 | -.397 | -.082 | .093 | -.109 | -.223 | .082 |
| | Blast Furnace Slag (kg/m³) | .135 | -.275 | 1.000 | -.324 | .107 | .043 | -.284 | -.282 | -.044 |
| | Fly Ash (kg/m³) | -.106 | -.397 | -.324 | 1.000 | -.257 | .377 | -.010 | .079 | -.154 |
| | Water (kg/m³) | -.290 | -.082 | .107 | -.257 | 1.000 | -.657 | -.182 | -.451 | .278 |
| | Superplasticizer (kg/m³) | .366 | .093 | .043 | .377 | -.657 | 1.000 | -.266 | .223 | -.193 |
| | Coarse Aggregate (kg/m³) | -.165 | -.109 | -.284 | -.010 | -.182 | -.266 | 1.000 | -.179 | -.003 |
| | Fine Aggregate (kg/m³) | -.167 | -.223 | -.282 | .079 | -.451 | .223 | -.179 | 1.000 | -.156 |
| | Curing Age (days) | .329 | .082 | -.044 | -.154 | .278 | -.193 | -.003 | -.156 | 1.000 |
| Sig. (1-tailed) | Compressive Strength (MPa) | . | <.001 | <.001 | <.001 | <.001 | <.001 | <.001 | <.001 | <.001 |
| | Cement (kg/m³) | .000 | . | .000 | .000 | .004 | .001 | .000 | .000 | .004 |
| | Blast Furnace Slag (kg/m³) | .000 | .000 | . | .000 | .000 | .082 | .000 | .000 | .078 |
| | Fly Ash (kg/m³) | .000 | .000 | .000 | . | .000 | .000 | .375 | .006 | .000 |
| | Water (kg/m³) | .000 | .004 | .000 | .000 | . | .000 | .000 | .000 | .000 |
| | Superplasticizer (kg/m³) | .000 | .001 | .082 | .000 | .000 | . | .000 | .000 | .000 |
| | Coarse Aggregate (kg/m³) | .000 | .000 | .000 | .375 | .000 | .000 | . | .000 | .461 |
| | Fine Aggregate (kg/m³) | .000 | .000 | .000 | .006 | .000 | .000 | .000 | . | .000 |
| | Curing Age (days) | .000 | .004 | .078 | .000 | .000 | .000 | .461 | .000 | . |
| N | Compressive Strength (MPa) | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 |
| | Cement (kg/m³) | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 |
| | Blast Furnace Slag (kg/m³) | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 |
| | Fly Ash (kg/m³) | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 |
| | Water (kg/m³) | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 |
| | Superplasticizer (kg/m³) | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 |
| | Coarse Aggregate (kg/m³) | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 |
| | Fine Aggregate (kg/m³) | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 |
| | Curing Age (days) | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 | 1030 |

Table 15: Correlation matrix between compressive strength and predictors.

**Variables Entered/Removed[a]**

| Model | Variables Entered | Variables Removed | Method |
|---|---|---|---|
| 1 | Cement (kg/m³) | . | Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100). |
| 2 | Superplasticizer (kg/m³) | . | Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100). |
| 3 | Curing Age (days) | . | Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100). |
| 4 | Blast Furnace Slag (kg/m³) | . | Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100). |
| 5 | Water (kg/m³) | . | Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100). |
| 6 | Fly Ash (kg/m³) | . | Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100). |

a. Dependent Variable: Compressive Strength (MPa)

Table 16: Stepwise entry and removal of predictors.

**Model Summary[g]**

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
|---|---|---|---|---|
| 1 | .498[a] | .248 | .247 | 14.495430857581399 |
| 2 | .593[b] | .351 | .350 | 13.470522571358584 |
| 3 | .694[c] | .482 | .480 | 12.045095584476996 |
| 4 | .742[d] | .551 | .549 | 11.216192566254911 |
| 5 | .764[e] | .584 | .582 | 10.798373985681206 |
| 6 | .784[f] | .614 | .612 | 10.409784091934055 |

a. Predictors: (Constant), Cement (kg/m³)

b. Predictors: (Constant), Cement (kg/m³), Superplasticizer (kg/m³)

c. Predictors: (Constant), Cement (kg/m³), Superplasticizer (kg/m³), Curing Age (days)

d. Predictors: (Constant), Cement (kg/m³), Superplasticizer (kg/m³), Curing Age (days), Blast Furnace Slag (kg/m³)

e. Predictors: (Constant), Cement (kg/m³), Superplasticizer (kg/m³), Curing Age (days), Blast Furnace Slag (kg/m³), Water (kg/m³)

f. Predictors: (Constant), Cement (kg/m³), Superplasticizer (kg/m³), Curing Age (days), Blast Furnace Slag (kg/m³), Water (kg/m³), Fly Ash (kg/m³)

g. Dependent Variable: Compressive Strength (MPa)

Table 17: Model summary for stepwise regression (Models 1–6).

**ANOVA[a]**

| Model | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| 1 | Regression | 71172.222 | 1 | 71172.222 | 338.726 | <.001[b] |
| | Residual | 216000.806 | 1028 | 210.118 | | |
| | Total | 287173.028 | 1029 | | | |
| 2 | Regression | 100818.766 | 2 | 50409.383 | 277.807 | <.001[c] |
| | Residual | 186354.263 | 1027 | 181.455 | | |
| | Total | 287173.028 | 1029 | | | |
| 3 | Regression | 138316.508 | 3 | 46105.503 | 317.784 | <.001[d] |
| | Residual | 148856.520 | 1026 | 145.084 | | |
| | Total | 287173.028 | 1029 | | | |
| 4 | Regression | 158224.978 | 4 | 39556.245 | 314.430 | <.001[e] |
| | Residual | 128948.050 | 1025 | 125.803 | | |
| | Total | 287173.028 | 1029 | | | |
| 5 | Regression | 167769.631 | 5 | 33553.926 | 287.757 | <.001[f] |
| | Residual | 119403.398 | 1024 | 116.605 | | |
| | Total | 287173.028 | 1029 | | | |
| 6 | Regression | 176317.061 | 6 | 29386.177 | 271.181 | <.001[g] |
| | Residual | 110855.968 | 1023 | 108.364 | | |
| | Total | 287173.028 | 1029 | | | |

a. Dependent Variable: Compressive Strength (MPa)

b. Predictors: (Constant), Cement (kg/m³)

c. Predictors: (Constant), Cement (kg/m³), Superplasticizer (kg/m³)

d. Predictors: (Constant), Cement (kg/m³), Superplasticizer (kg/m³), Curing Age (days)

e. Predictors: (Constant), Cement (kg/m³), Superplasticizer (kg/m³), Curing Age (days), Blast Furnace Slag (kg/m³)

f. Predictors: (Constant), Cement (kg/m³), Superplasticizer (kg/m³), Curing Age (days), Blast Furnace Slag (kg/m³), Water (kg/m³)

g. Predictors: (Constant), Cement (kg/m³), Superplasticizer (kg/m³), Curing Age (days), Blast Furnace Slag (kg/m³), Water (kg/m³), Fly Ash (kg/m³)

Table 18: ANOVA results for stepwise regression.

**Coefficients[a]**

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. | 95.0% Confidence Interval for B | | Collinearity Statistics | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | B | Std. Error | Beta | | | Lower Bound | Upper Bound | Tolerance | VIF |
| 1 | (Constant) | 13.443 | 1.297 | | 10.365 | <.001 | 10.898 | 15.988 | | |
| | Cement (kg/m³) | .080 | .004 | .498 | 18.405 | <.001 | .071 | .088 | 1.000 | 1.000 |
| 2 | (Constant) | 9.190 | 1.250 | | 7.350 | <.001 | 6.737 | 11.644 | | |
| | Cement (kg/m³) | .075 | .004 | .468 | 18.534 | <.001 | .067 | .083 | .991 | 1.009 |
| | Superplasticizer (kg/m³) | .902 | .071 | .323 | 12.782 | <.001 | .764 | 1.041 | .991 | 1.009 |
| 3 | (Constant) | 5.098 | 1.147 | | 4.446 | <.001 | 2.848 | 7.348 | | |
| | Cement (kg/m³) | .069 | .004 | .431 | 18.976 | <.001 | .062 | .076 | .981 | 1.019 |
| | Superplasticizer (kg/m³) | 1.112 | .064 | .397 | 17.246 | <.001 | .985 | 1.238 | .951 | 1.052 |
| | Curing Age (days) | .098 | .006 | .370 | 16.077 | <.001 | .086 | .110 | .953 | 1.050 |
| 4 | (Constant) | -2.001 | 1.208 | | -1.657 | .098 | -4.371 | .369 | | |
| | Cement (kg/m³) | .081 | .004 | .508 | 23.076 | <.001 | .074 | .088 | .905 | 1.105 |
| | Superplasticizer (kg/m³) | 1.060 | .060 | .379 | 17.610 | <.001 | .941 | 1.178 | .946 | 1.057 |
| | Curing Age (days) | .098 | .006 | .372 | 17.368 | <.001 | .087 | .110 | .953 | 1.050 |
| | Blast Furnace Slag (kg/m³) | .053 | .004 | .275 | 12.580 | <.001 | .045 | .061 | .919 | 1.088 |
| 5 | (Constant) | 35.492 | 4.304 | | 8.246 | <.001 | 27.046 | 43.938 | | |
| | Cement (kg/m³) | .081 | .003 | .509 | 24.007 | <.001 | .075 | .088 | .905 | 1.105 |
| | Superplasticizer (kg/m³) | .614 | .076 | .219 | 8.073 | <.001 | .465 | .763 | .549 | 1.821 |
| | Curing Age (days) | .109 | .006 | .413 | 19.554 | <.001 | .098 | .120 | .910 | 1.100 |
| | Blast Furnace Slag (kg/m³) | .060 | .004 | .311 | 14.520 | <.001 | .052 | .068 | .888 | 1.127 |
| | Water (kg/m³) | -.197 | .022 | -.252 | -9.047 | <.001 | -.240 | -.154 | .524 | 1.908 |
| 6 | (Constant) | 29.030 | 4.212 | | 6.891 | <.001 | 20.764 | 37.296 | | |
| | Cement (kg/m³) | .105 | .004 | .660 | 24.821 | <.001 | .097 | .114 | .534 | 1.871 |
| | Superplasticizer (kg/m³) | .239 | .085 | .085 | 2.826 | .005 | .073 | .405 | .412 | 2.424 |
| | Curing Age (days) | .113 | .005 | .429 | 20.987 | <.001 | .103 | .124 | .902 | 1.108 |
| | Blast Furnace Slag (kg/m³) | .086 | .005 | .447 | 17.386 | <.001 | .077 | .096 | .572 | 1.749 |
| | Water (kg/m³) | -.218 | .021 | -.279 | -10.332 | <.001 | -.260 | -.177 | .517 | 1.933 |
| | Fly Ash (kg/m³) | .069 | .008 | .263 | 8.881 | <.001 | .054 | .084 | .430 | 2.328 |

a. Dependent Variable: Compressive Strength (MPa)

Table 19: Coefficients table for regression models (final model highlighted).

**Excluded Variables[a]**

| Model | | Beta In | t | Sig. | Partial Correlation | Collinearity Statistics Tolerance | VIF | Minimum Tolerance |
|---|---|---|---|---|---|---|---|---|
| 1 | Blast Furnace Slag (kg/m³) | .294[b] | 11.051 | <.001 | .326 | .924 | 1.082 | .924 |
| | Fly Ash (kg/m³) | .109[b] | 3.735 | <.001 | .116 | .842 | 1.188 | .842 |
| | Water (kg/m³) | -.251[b] | -9.641 | <.001 | -.288 | .993 | 1.007 | .993 |
| | Superplasticizer (kg/m³) | .323[b] | 12.782 | <.001 | .370 | .991 | 1.009 | .991 |
| | Coarse Aggregate (kg/m³) | -.112[b] | -4.141 | <.001 | -.128 | .988 | 1.012 | .988 |
| | Fine Aggregate (kg/m³) | -.059[b] | -2.141 | .032 | -.067 | .950 | 1.052 | .950 |
| | Curing Age (days) | .290[b] | 11.329 | <.001 | .333 | .993 | 1.007 | .993 |
| 2 | Blast Furnace Slag (kg/m³) | .271[c] | 10.936 | <.001 | .323 | .919 | 1.088 | .913 |
| | Fly Ash (kg/m³) | -.062[c] | -2.024 | .043 | -.063 | .669 | 1.495 | .669 |
| | Water (kg/m³) | -.069[c] | -2.079 | .038 | -.065 | .567 | 1.763 | .566 |
| | Coarse Aggregate (kg/m³) | -.030[c] | -1.153 | .249 | -.036 | .922 | 1.085 | .922 |
| | Fine Aggregate (kg/m³) | -.151[c] | -5.772 | <.001 | -.177 | .891 | 1.123 | .891 |
| | Curing Age (days) | .370[c] | 16.077 | <.001 | .449 | .953 | 1.050 | .951 |
| 3 | Blast Furnace Slag (kg/m³) | .275[d] | 12.580 | <.001 | .366 | .919 | 1.088 | .905 |
| | Fly Ash (kg/m³) | -.041[d] | -1.495 | .135 | -.047 | .667 | 1.498 | .667 |
| | Water (kg/m³) | -.177[d] | -5.887 | <.001 | -.181 | .543 | 1.842 | .543 |
| | Coarse Aggregate (kg/m³) | -.012[d] | -.504 | .614 | -.016 | .920 | 1.087 | .885 |
| | Fine Aggregate (kg/m³) | -.116[d] | -4.884 | <.001 | -.151 | .882 | 1.133 | .882 |
| 4 | Fly Ash (kg/m³) | .228[e] | 7.387 | <.001 | .225 | .435 | 2.298 | .435 |
| | Water (kg/m³) | -.252[e] | -9.047 | <.001 | -.272 | .524 | 1.908 | .524 |
| | Coarse Aggregate (kg/m³) | .086[e] | 3.739 | <.001 | .116 | .824 | 1.214 | .824 |
| | Fine Aggregate (kg/m³) | -.004[e] | -.167 | .868 | -.005 | .741 | 1.349 | .741 |
| 5 | Fly Ash (kg/m³) | .263[f] | 8.881 | <.001 | .268 | .430 | 2.328 | .412 |
| | Coarse Aggregate (kg/m³) | -.011[f] | -.455 | .650 | -.014 | .641 | 1.560 | .408 |
| | Fine Aggregate (kg/m³) | -.103[f] | -4.062 | <.001 | -.126 | .628 | 1.593 | .444 |
| 6 | Coarse Aggregate (kg/m³) | .016[g] | .639 | .523 | .020 | .631 | 1.584 | .346 |
| | Fine Aggregate (kg/m³) | .015[g] | .521 | .603 | .016 | .457 | 2.186 | .313 |

a. Dependent Variable: Compressive Strength (MPa)

b. Predictors in the Model: (Constant), Cement (kg/m³)

c. Predictors in the Model: (Constant), Cement (kg/m³), Superplasticizer (kg/m³)

d. Predictors in the Model: (Constant), Cement (kg/m³), Superplasticizer (kg/m³), Curing Age (days)

e. Predictors in the Model: (Constant), Cement (kg/m³), Superplasticizer (kg/m³), Curing Age (days), Blast Furnace Slag (kg/m³)

f. Predictors in the Model: (Constant), Cement (kg/m³), Superplasticizer (kg/m³), Curing Age (days), Blast Furnace Slag (kg/m³), Water (kg/m³)

g. Predictors in the Model: (Constant), Cement (kg/m³), Superplasticizer (kg/m³), Curing Age (days), Blast Furnace Slag (kg/m³), Water (kg/m³), Fly Ash (kg/m³)

Table 20: Variables excluded during stepwise regression.

**Collinearity Diagnostics[a]**

| Model | Dimension | Eigenvalue | Condition Index | Variance Proportions (Constant) | Cement (kg/m³) | Superplasticizer (kg/m³) | Curing Age (days) | Blast Furnace Slag (kg/m³) | Water (kg/m³) | Fly Ash (kg/m³) |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1.937 | 1.000 | .03 | .03 | | | | | |
| | 2 | .063 | 5.563 | .97 | .97 | | | | | |
| 2 | 1 | 2.576 | 1.000 | .02 | .02 | .05 | | | | |
| | 2 | .362 | 2.666 | .04 | .06 | .93 | | | | |
| | 3 | .062 | 6.448 | .95 | .93 | .01 | | | | |
| 3 | 1 | 2.956 | 1.000 | .01 | .01 | .03 | .03 | | | |
| | 2 | .698 | 2.058 | .00 | .00 | .21 | .61 | | | |
| | 3 | .285 | 3.221 | .06 | .10 | .74 | .35 | | | |
| | 4 | .062 | 6.929 | .93 | .89 | .02 | .01 | | | |
| 4 | 1 | 3.401 | 1.000 | .01 | .01 | .02 | .02 | .02 | | |
| | 2 | .717 | 2.178 | .00 | .00 | .13 | .62 | .07 | | |
| | 3 | .549 | 2.489 | .00 | .01 | .16 | .01 | .70 | | |
| | 4 | .282 | 3.474 | .04 | .11 | .68 | .34 | .02 | | |
| | 5 | .051 | 8.181 | .95 | .87 | .01 | .01 | .19 | | |
| 5 | 1 | 4.325 | 1.000 | .00 | .01 | .01 | .01 | .01 | .00 | |
| | 2 | .718 | 2.454 | .00 | .00 | .08 | .57 | .07 | .00 | |
| | 3 | .549 | 2.806 | .00 | .01 | .09 | .01 | .68 | .00 | |
| | 4 | .332 | 3.611 | .00 | .04 | .36 | .37 | .03 | .00 | |
| | 5 | .073 | 7.717 | .01 | .93 | .01 | .00 | .19 | .02 | |
| | 6 | .003 | 36.016 | .98 | .01 | .45 | .03 | .01 | .98 | |
| 6 | 1 | 4.736 | 1.000 | .00 | .00 | .01 | .01 | .01 | .00 | .01 |
| | 2 | .880 | 2.320 | .00 | .00 | .04 | .18 | .05 | .00 | .12 |
| | 3 | .682 | 2.636 | .00 | .00 | .01 | .35 | .24 | .00 | .02 |
| | 4 | .361 | 3.622 | .00 | .06 | .01 | .31 | .19 | .00 | .12 |
| | 5 | .305 | 3.941 | .00 | .00 | .40 | .11 | .02 | .00 | .14 |
| | 6 | .033 | 11.995 | .03 | .92 | .23 | .01 | .49 | .04 | .59 |
| | 7 | .003 | 37.721 | .97 | .01 | .31 | .03 | .00 | .95 | .00 |

a. Dependent Variable: Compressive Strength (MPa)

Table 21: Collinearity diagnostics for stepwise regression.

**Residuals Statistics[a]**

| | Minimum | Maximum | Mean | Std. Deviation | N |
|---|---|---|---|---|---|
| Predicted Value | 10.41257762908936 | 78.84008026123047 | 35.81783582611354 | 13.089995019369182 | 1030 |
| Residual | -29.013690948486328 | 34.726081848144530 | .000000000000004 | 10.379390496869727 | 1030 |
| Std. Predicted Value | -1.941 | 3.287 | .000 | 1.000 | 1030 |
| Std. Residual | -2.787 | 3.336 | .000 | .997 | 1030 |

a. Dependent Variable: Compressive Strength (MPa)

Figure 6: Residuals statistics for the regression model.

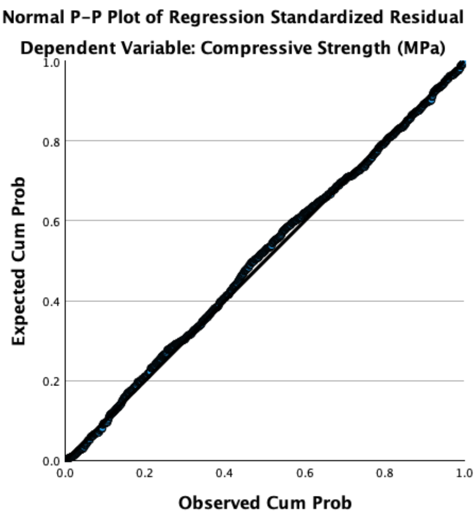


Figure 7: Normal P–P plot of regression standardized residuals.

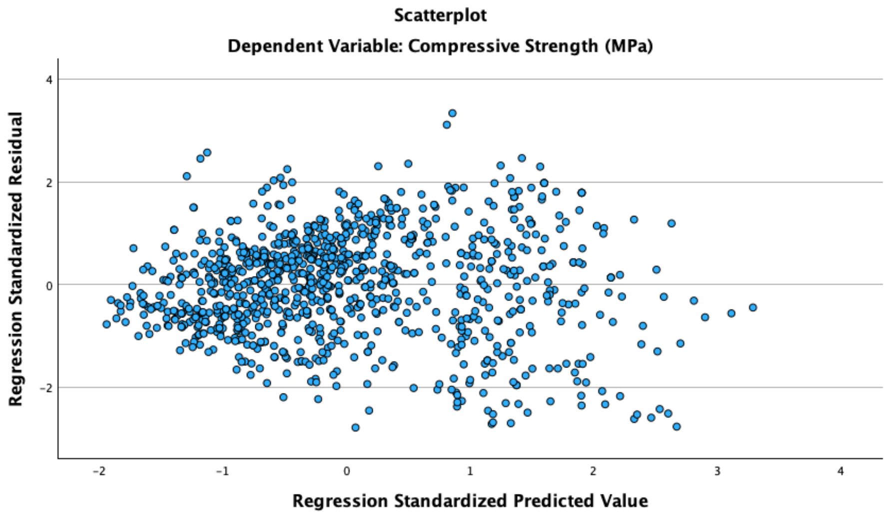


Figure 8: Scatterplot of standardized residuals versus predicted values.

**Case Processing Summary**

| | Cases | | | | | |
|---|---|---|---|---|---|---|
| | Valid | | Missing | | Total | |
| | N | Percent | N | Percent | N | Percent |
| Unstandardized Residual | 1030 | 100.0% | 0 | 0.0% | 1030 | 100.0% |

Table 22: Case processing summary for residuals.

**Descriptives**

| | | | Statistic | Std. Error |
|---|---|---|---|---|
| Unstandardized Residual | Mean | | .0000000 | .32340985 |
| | 95% Confidence Interval for Mean | Lower Bound | -.6346181 | |
| | | Upper Bound | .6346181 | |
| | 5% Trimmed Mean | | .1328851 | |
| | Median | | .6501919 | |
| | Variance | | 107.732 | |
| | Std. Deviation | | 10.37939050 | |
| | Minimum | | -29.01369 | |
| | Maximum | | 34.72608 | |
| | Range | | 63.73977 | |
| | Interquartile Range | | 13.03322 | |
| | Skewness | | -.174 | .076 |
| | Kurtosis | | .025 | .152 |

Table 23: Descriptive statistics of unstandardized residuals.

**Tests of Normality**

| | Kolmogorov-Smirnov[a] | | | Shapiro-Wilk | | |
|---|---|---|---|---|---|---|
| | Statistic | df | Sig. | Statistic | df | Sig. |
| Unstandardized Residual | .031 | 1030 | .025 | .996 | 1030 | .007 |

a. Lilliefors Significance Correction

Table 24: Kolmogorov–Smirnov and Shapiro–Wilk tests of normality.

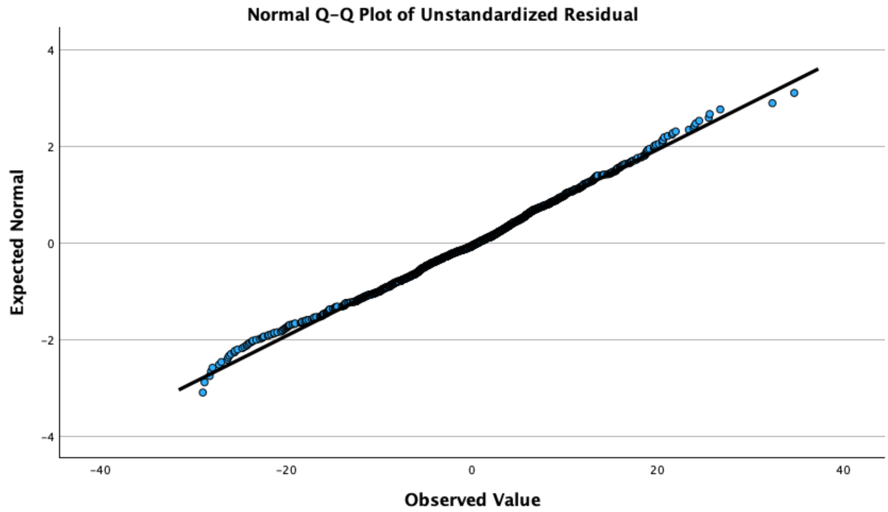


Figure 9: Normal Q–Q plot of residuals.

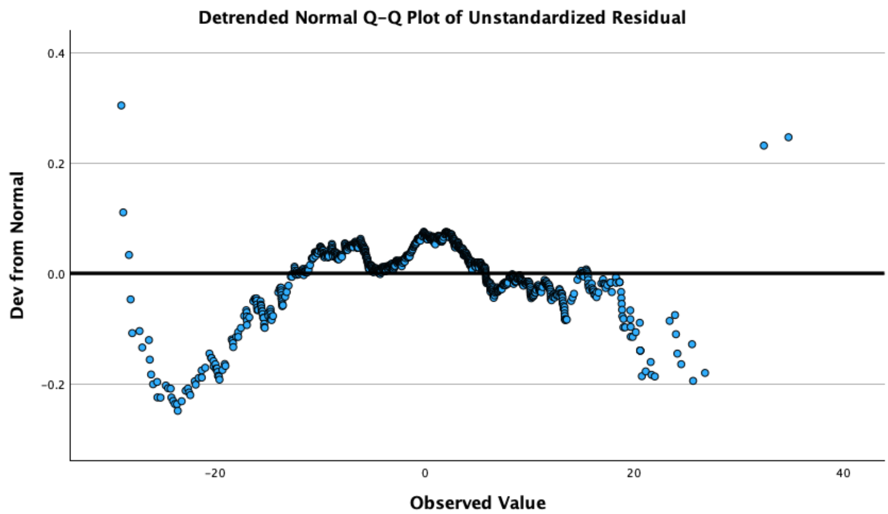


Figure 10: Detrended Q–Q plot of residuals.

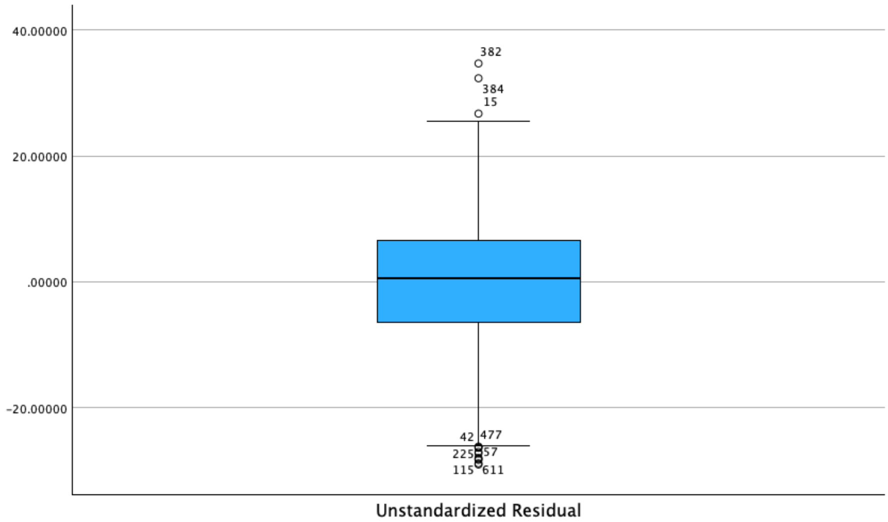


Figure 11: Boxplot of unstandardized residuals.