

Concrete Composition & Strength Analysis

Introduction

This project analyzes a concrete manufacturing dataset containing **1030 concrete samples**, each with measured **ratios of cementitious and aggregate materials** relative to water, age (in days), and compressive strength. For every mix, we observe:

- cement / water ratio
- slag / water ratio
- fly ash / water ratio
- superplasticizer / water ratio
- coarse aggregate / water ratio
- fine aggregate / water ratio
- age
- compressive strength (MPa)

The analysis utilizes **Analysis of Variance (ANOVA)**, **Multiple Linear Regression**, and **Linear Discriminant Analysis (LDA)** to address the client's specific questions.

Overview

Concrete strength is known to vary both with mix composition and curing age, but the magnitude, shape, and interaction of these effects are not always straightforward. The goal of this project is to apply statistical methods to answer five key analysis questions:

1. How do strength levels vary across compositions and age groups?
2. Can we identify natural clusters of concrete mixes, and do they differ in strength?
3. Can we predict compressive strength for concrete aged =100 days?
4. Can we predict whether concrete aged 90–100 days will exceed 50 MPa?
5. Can mixture composition and strength reliably classify concrete into age groups?

Section 1 — Descriptive Exploration of Strength, Composition, and Age

1.1 Summary of Compressive Strength

We examined *key summary statistics* to understand the overall strength behavior in the dataset. Concrete strength shows **considerable spread and skewness**, indicating different mixture compositions, curing durations, and material contributions.

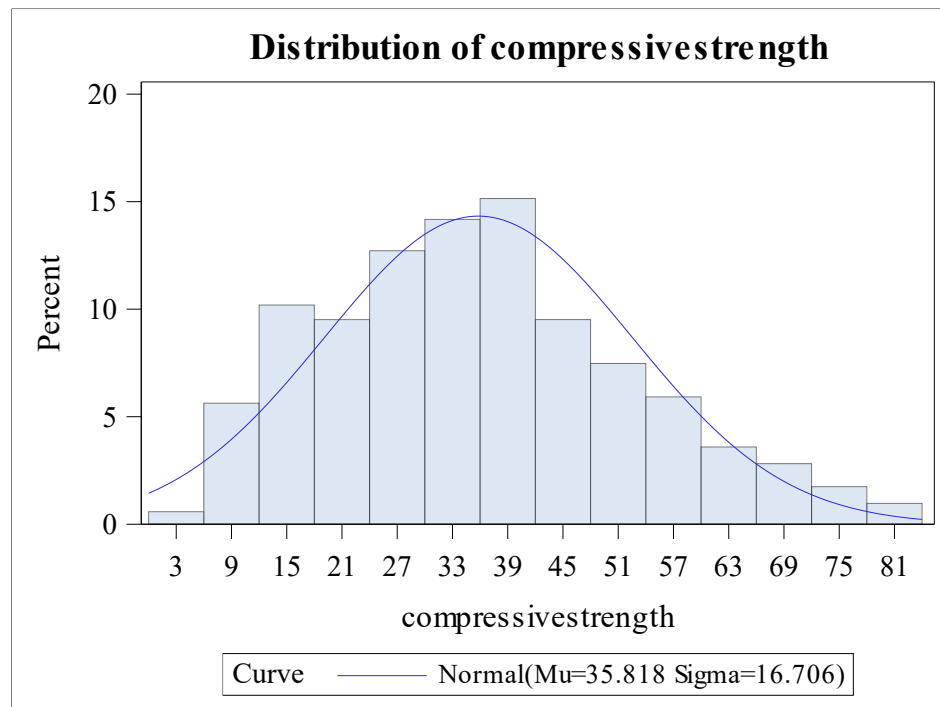
- Minimum: 2.3 MPa
- Maximum: 82.6 MPa
- Mean: 35.8 MPa
- Median: 33.6 MPa
- Standard deviation: 16.7 MPa
- Distribution type: Right-skewed
- Most values lie between 20–50 MPa

Variable	N	Mean	Std Dev	Minimum	Median	Maximum
cementwater	1030	1.5782748	0.6481051	0.5312500	1.4807176	3.7468265
slagwater	1030	0.4068528	0.4719605	0	0.1273885	1.9353796
flyashwater	1030	0.3134171	0.3756394	0	0	1.3456263
superplasticizerwater	1030	0.0374018	0.0391309	0	0.0346633	0.2336720
coarsewater	1030	5.4431809	0.8429658	3.4534413	5.4518040	8.6956879
finewater	1030	4.3447628	0.8249082	2.6052632	4.2994792	7.8404423
age	1030	45.6621359	63.1699116	1.0000000	28.0000000	365.0000000
compressivestrength	1030	35.8178358	16.7056792	2.3318078	34.4427736	82.5992248

1.2 Histogram of Compressive Strength

The histogram shows that **most mixes fall** in the *mid-strength range*, while a **smaller number** reach very *low or very high strengths*. This visual confirms the right-skewed nature observed in the statistics.

- Main cluster: 20–50 MPa
- High-strength tail extends to 82.6 MPa
- Low-strength mixes visible near 5–15 MPa
- Skew: Right-skewed distribution



1.3 Correlation with Mix Ratios & Age

Correlation analysis highlights which materials *contribute positively or negatively to strength* development. **Cement and superplasticizer** ratios show **strong associations**, while **fly ash and fine aggregate** exhibit **mild negative** relationships.

- Cement/water: $r \sim 0.56$
- Superplasticizer/water: $r \sim 0.38$
- Age: $r \sim 0.33$
- Slag/water: $r \sim 0.15-0.20$
- Fly ash/water: $r \sim -0.10$
- Coarse aggregate: $r \sim 0.05$
- Fine aggregate: $r \sim -0.12$

7 with Variables:	Cementwater slagwater flyashwater superplasticizerwater coarsewater finewater age
1 Variables:	compressivestrength

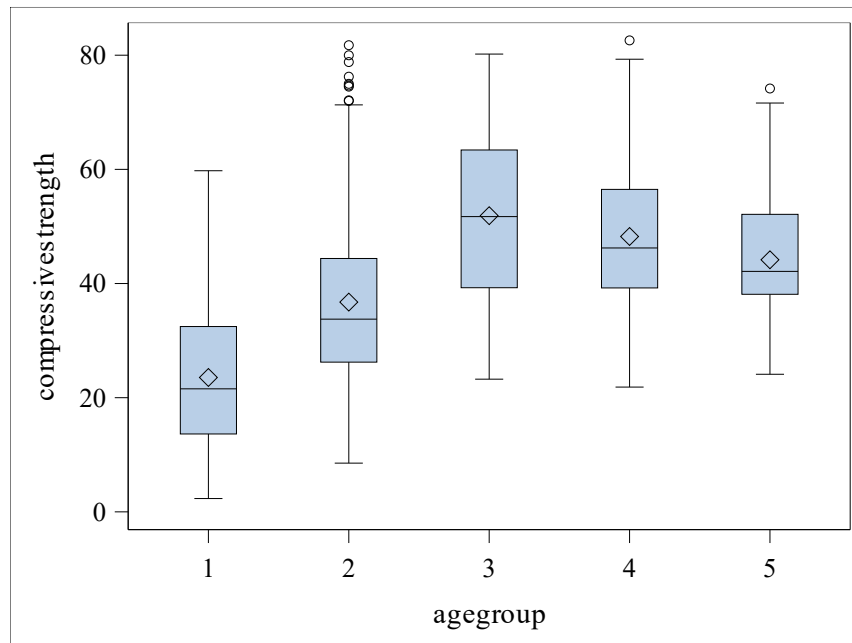
Pearson Correlation Coefficients, N = 1030 Prob > r under H0: Rho=0	
	compressivestrength
cementwater	0.55952 <.0001
slagwater	0.18206 <.0001
flyashwater	-0.08597 0.0058
superplasticizerwater	0.37867 <.0001
coarsewater	0.15853 <.0001
finewater	0.12707 <.0001
age	0.32888 <.0001

1.4 Strength Across Age Groups

Strength was compared across *five curing age groups* to evaluate hydration effects. Strength **increases** sharply in **young concrete** and then **stabilizes** after **two to three months** as hydration reactions reach completion.

- Group 1 (<28 days): ~23 MPa
- Group 2 (28–55 days): ~36 MPa
- Group 3 (56–89 days): ~52 MPa
- Group 4 (90–179 days): ~48 MPa
- Group 5 (=180 days): ~44 MPa
- Strength plateaus after ~90 days

Analysis Variable : compressivestrength						
agegroup	N Obs	Mean	Median	Std Dev	Minimum	Maximum
1	324	23.54	21.55	12.41	2.33	59.76
2	425	36.75	33.76	14.71	8.54	81.75
3	91	51.89	51.72	14.31	23.25	80.20
4	131	48.24	46.23	13.50	21.86	82.60
5	59	44.16	42.13	10.61	24.10	74.17



Section 2 — Cluster Analysis of Concrete Mixes

2.1 Standardization & Ward's Method

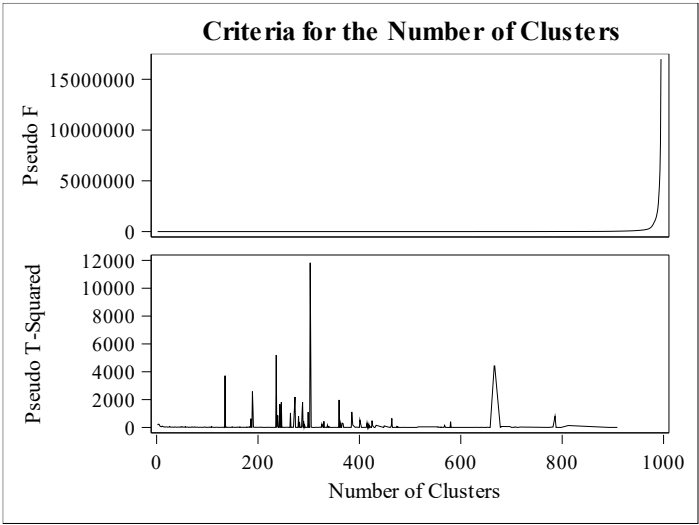
Clustering was applied to identify natural groupings of concrete mixtures based on *similarity in component ratios and age*. Variables were standardized to ensure equal importance in distance calculations.

- Variables standardized: 7
- Method used: Ward's hierarchical clustering
- Selected clusters: 4
- Purpose: identify distinct mixture families

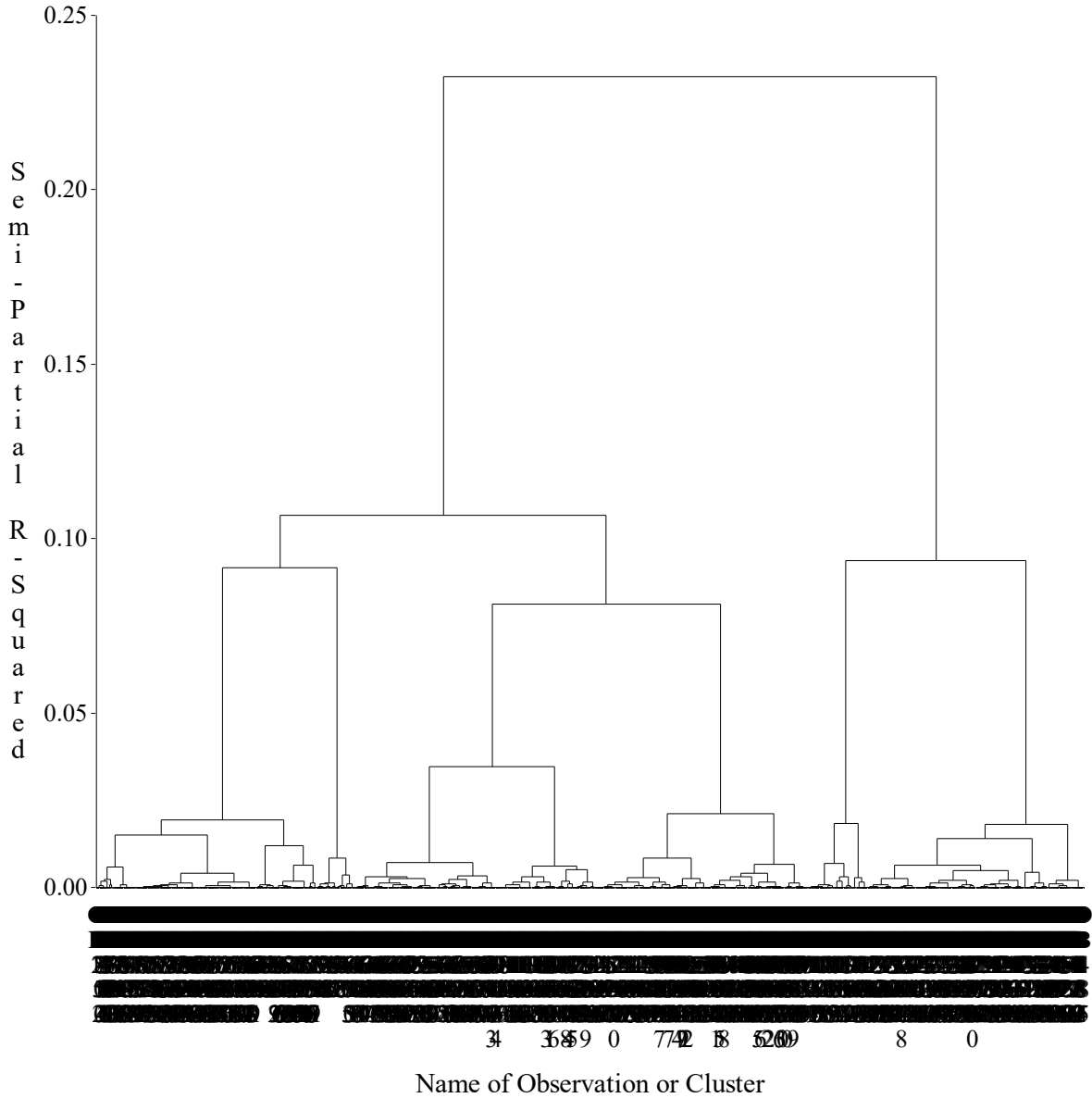
Eigenvalues of the Covariance Matrix				
	Eigenvalue	Difference	Proportion	Cumulative
1	2.57065256	1.23205210	0.3672	0.3672
2	1.33860045	0.17351463	0.1912	0.5585
3	1.16508582	0.32195833	0.1664	0.7249
4	0.84312750	0.24480057	0.1204	0.8454
5	0.59832693	0.24586780	0.0855	0.9308
6	0.35245913	0.22071152	0.0504	0.9812
7	0.13174761		0.0188	1.0000

Root-Mean-Square Total-Sample Standard Deviation	1
--	---

Root-Mean-Square Distance Between Observations	3.741657
--	----------



Ward's Minimum Variance Cluster Analysis



2.2 Cluster Profiles

Clusters represent different mixture strategies with distinct strength outcomes. Understanding these profiles helps identify which material combinations tend to **produce stronger or weaker concretes**.

- Cluster 1: ~33 MPa, high slag + fly ash
- Cluster 2: ~54 MPa, highest cement + superplasticizer
- Cluster 3: ~31 MPa, balanced mixes, lowest strength
- Cluster 4: ~44 MPa, very old mixes (~260 days)

CLUSTER	N Obs	Variable	Mean	Median	Std Dev
1	63	compressivestrength	55.20	56.70	14.45
		age	-0.14	-0.28	0.51
		cementwater	1.84	1.84	0.67
		slagwater	0.81	0.63	0.72
		flyashwater	-0.83	-0.83	0.00
		superplasticizerwater	2.25	1.79	1.31
		r	0.99	0.88	0.81
		coarsewater finewater	1.74	1.74	1.03
2	470	compressivestrength	34.28	32.86	16.97
		age	-0.22	-0.28	0.43
		cementwater	-0.46	-0.62	0.79
		slagwater	0.63	0.72	0.99
		flyashwater	-0.03	-0.65	0.89
		superplasticizerwater	-0.09	-0.06	0.67
		r	-0.48	-0.50	0.68
		coarsewater finewater	-0.33	-0.29	0.62
3	227	compressivestrength	36.17	36.80	14.46
		age	-0.15	-0.28	0.51
		cementwater	-0.06	-0.30	0.80
		slagwater	-0.68	-0.86	0.29
		flyashwater	1.27	1.19	0.54
		superplasticizerwater	0.62	0.63	0.55
		r	1.09	0.93	0.68
		coarsewater finewater	0.87	0.81	0.66
4	270	compressivestrength	33.67	32.87	15.65
		age	0.54	-0.28	1.68
		cementwater	0.42	0.22	0.86
		slagwater	-0.71	-0.86	0.43
		flyashwater	-0.83	-0.83	0.08
		superplasticizerwater	-0.89	-0.96	0.26
		r	-0.30	-0.41	0.87
		coarsewater finewater	-0.56	-0.32	0.84

2.3 ANOVA: Strength Differences Across Clusters

ANOVA was used to evaluate strength differences between clusters. The *strong significance* supports clear distinctions in mix **behavior and performance**.

- ANOVA p-value: <0.0001
- Strongest: Cluster 2 (~54 MPa)
- Weakest: Cluster 3 (~31 MPa)
- Strength difference across clusters: ~23 MPa

Class Level Information		
Class	Levels	Values
CLUSTER	4	1 2 3 4

Dependent Variable: compressivestrength

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	26056.1087	8685.3696	34.13	<.0001
Error	1026	261116.9198	254.4999		
Corrected Total	1029	287173.0285			

R-Square	Coeff Var	Root MSE	compressivestrength Mean
0.090733	44.53941	15.95305	35.81784

Source	DF	Type I SS	Mean Square	F Value	Pr > F
CLUSTER	3	26056.10869	8685.36956	34.13	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
CLUSTER	3	26056.10869	8685.36956	34.13	<.0001

Levene's Test for Homogeneity of compressivestrength Variance ANOVA of Squared Deviations from Group Means					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
CLUSTER	3	1173377	391126	3.31	0.0195
Error	1026	1.2113E8	118061		

Tukey's Studentized Range (HSD) Test for compressivestrength

Note: This test controls the Type I experimentwise error rate.

Alpha	0.05
Error Degrees of Freedom	1026
Error Mean Square	254.4999
Critical Value of Studentized Range	3.63906

Comparisons significant at the 0.05 level are indicated by ***.				
CLUSTER Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
1 - 3	19.038	13.193	24.884	***
1 - 2	20.921	15.414	26.429	***
1 - 4	21.533	15.789	27.277	***
3 - 1	-19.038	-24.884	-13.193	***
3 - 2	1.883	-1.435	5.201	
3 - 4	2.495	-1.202	6.191	
2 - 1	-20.921	-26.429	-15.414	***
2 - 3	-1.883	-5.201	1.435	
2 - 4	0.612	-2.523	3.747	
4 - 1	-21.533	-27.277	-15.789	***
4 - 3	-2.495	-6.191	1.202	
4 - 2	-0.612	-3.747	2.523	

Section 3 — Regression Model for Strength (Age =100 Days)

3.1 Objective

A regression model was built for **mature concrete (=100 days)** to isolate the effects of mix components on *long-term strength*, independent of early hydration.

- Sample size: ~114 mature samples
- Goal: predict strength purely from composition
- Removes early-age noise

3.2 Regression Results

Stepwise regression identified key mix components that significantly influence **long-term strength**.

Cement, slag, and superplasticizer improve strength, while **excessive fine aggregate has a slight weakening effect**.

- Predictors selected:
 - Cement/water (+14.5 MPa)
 - Slag/water (+15 MPa)
 - Superplasticizer/water (+260 MPa)
 - Fine aggregate/water (–2.6 MPa)
- *Model R²: 0.56
- *Age not included ? strength plateau

Model: MODEL1
Dependent Variable: compressivestrength

Stepwise Selection: Step 1
Variable superplasticizerwater Entered: R-Square = 0.1236 and C(p) = 108.1763

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	1307.65068	1307.65068	15.80	0.0001
Error	112	9270.00408	82.76789		
Corrected Total	113	10578			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	42.97342	1.08487	129870	1569.08	<.0001
superplasticizerwater	120.67287	30.35951	1307.65068	15.80	0.0001

Bounds on condition number: 1, 1

Stepwise Selection: Step 5
Variable coarsewater Entered: R-Square = 0.5683 and C(p) = 5.4759

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	6011.15357	1202.23071	28.43	<.0001
Error	108	4566.50119	42.28242		
Corrected Total	113	10578			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	24.93021	5.92829	747.74348	17.68	<.0001
cementwater	12.37448	2.16622	1379.77449	32.63	<.0001
slagwater	13.14266	2.87261	885.05885	20.93	<.0001
superplasticizerwater	230.78353	37.06693	1639.06499	38.76	<.0001
coarsewater	2.57531	1.52737	120.20630	2.84	0.0947
finewater	-4.64484	1.61913	347.96775	8.23	0.0050

Bounds on condition number: 6.6736, 102.15

All variables left in the model are significant at the 0.1500 level.

No other variable met the 0.1500 significance level for entry into the model.

Summary of Stepwise Selection								
Step	Variable Entered	Variable Removed	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	superplasticizerwater		1	0.1236	0.1236	108.176	15.80	0.0001
2	cementwater		2	0.1787	0.3024	65.6789	28.44	<.0001
3	slagwater		3	0.2311	0.5335	10.1366	54.50	<.0001
4	finewater		4	0.0234	0.5569	6.3051	5.76	0.0181
5	coarsewater		5	0.0114	0.5683	5.4759	2.84	0.0947

Model: MODEL1

Dependent Variable: compressivestrength

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	6011.15357	1202.23071	28.43	<.0001
Error	108	4566.50119	42.28242		
Corrected Total	113	10578			

Section 4 — Logistic Model: Predicting =50 MPa Strength (Age 90–100 Days)

4.1 Setup

A logistic regression model was used to *classify* whether concrete **aged 90–100 days** achieves at least **50 MPa**, a common structural threshold.

- Window analyzed: 90–100 days
- Binary target: 1 = =50 MPa
- Sample size: ~128 mixes

4.2 Logistic Results

The *model performed very well*, with **cement, slag, and superplasticizer increasing strength likelihood**, and coarse **aggregate and fly ash decreasing it**.

- Predictors selected: cement, slag, superplasticizer, fly ash, coarse aggregate
- Accuracy: ~95%
- AUC: 0.90–0.93
- Misclassification: <5%

Model Information	
Data Set	WORK.AGE90100
Response Variable	strong50
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	128
Number of Observations Used	128

Response Profile		
Ordered Value	strong50	Total Frequency
1	1	51
2	0	77

Probability modeled is strong50=1.

Stepwise Selection Procedure

Step 0. Intercept entered:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

-2 Log L	=	172.127
-----------------	---	---------

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
74.3184	6	<.0001

Step 1. Effect superplasticizerwate entered:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	174.127	126.767
SC	176.980	132.472
-2 Log L	172.127	122.767

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	49.3600	1	<.0001
Score	39.4759	1	<.0001
Wald	30.3747	1	<.0001

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
48.4933	5	<.0001

Note: No effects for the model in Step 1 are removed.

Step 7. Effect finewater is removed:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	174.127	64.926
SC	176.980	82.038
-2 Log L	172.127	52.926

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	119.2020	5	<.0001
Score	74.2576	5	<.0001
Wald	26.8363	5	<.0001

Residual Chi-Square Test		
Chi-Square	DF	Pr > ChiSq
0.5269	1	0.4679

Note: No effects for the model in Step 7 are removed.

Note: No (additional) effects met the 0.15 significance level for entry into the model.

Summary of Stepwise Selection							
Step	Effect		DF	Number In	Score Chi-Square	Wald Chi-Square	Pr > ChiSq
	Entered	Removed					
1	superplasticizerwate		1	1	39.4759		<.0001
2	cementwater		1	2	22.0435		<.0001
3	slagwater		1	3	26.6605		<.0001
4	finewater		1	4	4.2477		0.0393
5	flyashwater		1	5	5.3045		0.0213
6	coarsewater		1	6	4.2494		0.0393
7		finewater	1	5		0.5211	0.4704

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-7.3006	3.1820	5.2640	0.0218
cementwater	1	7.9146	1.6476	23.0768	<.0001
slagwater	1	8.1185	1.7508	21.5019	<.0001
flyashwater	1	6.8592	2.4441	7.8760	0.0050
superplasticizerwate	1	79.9769	24.8791	10.3338	0.0013
coarsewater	1	-2.3090	0.8019	8.2918	0.0040

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
cementwater	>999.999	108.358	>999.999
slagwater	>999.999	108.526	>999.999
flyashwater	952.570	7.915	>999.999
superplasticizerwate	>999.999	>999.999	>999.999
coarsewater	0.099	0.021	0.478

Section 5 — Discriminant Analysis for Age Group Classification

5.1 Purpose

Linear Discriminant Analysis (LDA) was applied to *classify samples into age groups* based on mix **composition and strength patterns**, highlighting how well curing stage can be inferred from material characteristics.

- Groups classified: 5
- Predictors used: 7 ratios + strength
- Method: LDA with cross-validation

Class Level Information					
agegroup	Variable Name	Frequency	Weight	Proportion	Prior Probability
1	_1	324	324.0000	0.314563	0.200000
2	_2	425	425.0000	0.412621	0.200000
3	_3	91	91.0000	0.088350	0.200000
4	_4	131	131.0000	0.127184	0.200000
5	_5	59	59.0000	0.057282	0.200000

Within Covariance Matrix Information		
agegroup	Covariance Matrix Rank	Natural Log of the Determinant of the Covariance Matrix
1	7	-10.30850
2	7	-9.99257
3	7	-11.58309
4	7	-11.24646
5	5	-51.09543
Pooled	7	-9.98953

Test of Homogeneity of Within Covariance Matrices

Chi-Square	DF	Pr > ChiSq
2725.661599	112	<.0001

Since the Chi-Square value is significant at the 0.1 level, the within covariance matrices will be used in the discriminant function.

Reference: Morrison, D.F. (1976) Multivariate Statistical Methods p252.

5.2 LDA Results

The model shows moderate accuracy overall. **Early-age** and **long-cured** concrete are easiest to classify, while **middle-age** groups **overlap** due to *similar strength ranges*.

- Overall accuracy: 65–70%
- Group 1: ~80%
- Group 2: ~78%
- Groups 3 & 4: ~30–40%
- Group 5: ~42%

Generalized Squared Distance to agegroup					
From agegroup	1	2	3	4	5
1	-10.30850	-5.60308	3.92801	6.87772	218161983
2	-5.88638	-9.99257	-6.86260	-6.08877	311840244
3	1.11968	-7.16998	-11.58309	-9.99132	706816600
4	2.17521	-7.20291	-9.89684	-11.24646	228176541
5	3.62571	-4.96030	-2.20187	-9.35542	-51.09543

Classification Summary for Calibration Data: WORK.CONCRETERATIOS *Resubstitution Summary using Quadratic Discriminant Function*

Number of Observations and Percent Classified into agegroup						
From agegroup	1	2	3	4	5	Total
1	162 50.00	27 8.33	18 5.56	0 0.00	117 36.11	324 100.00
2	27 6.35	223 52.47	74 17.41	24 5.65	77 18.12	425 100.00
3	1 1.10	4 4.40	68 74.73	16 17.58	2 2.20	91 100.00
4	0 0.00	3 2.29	45 34.35	30 22.90	53 40.46	131 100.00
5	0 0.00	0 0.00	0 0.00	0 0.00	59 100.00	59 100.00
Total	190 18.45	257 24.95	205 19.90	70 6.80	308 29.90	1030 100.00
Priors	0.2	0.2	0.2	0.2	0.2	

Error Count Estimates for agegroup						
	1	2	3	4	5	Total
Rate	0.5000	0.4753	0.2527	0.7710	0.0000	0.3998
Priors	0.2000	0.2000	0.2000	0.2000	0.2000	

Classification Summary for Calibration Data: WORK.CONCRETERATIOS

Cross-validation Summary using Quadratic Discriminant Function

Number of Observations and Percent Classified into agegroup						
From agegroup	1	2	3	4	5	Total
1	160 49.38	27 8.33	20 6.17	0 0.00	117 36.11	324 100.00
2	29 6.82	221 52.00	74 17.41	24 5.65	77 18.12	425 100.00
3	1 1.10	7 7.69	57 62.64	24 26.37	2 2.20	91 100.00
4	0 0.00	3 2.29	52 39.69	23 17.56	53 40.46	131 100.00
5	0 0.00	0 0.00	0 0.00	0 0.00	59 100.00	59 100.00
Total	190 18.45	258 25.05	203 19.71	71 6.89	308 29.90	1030 100.00
Priors	0.2	0.2	0.2	0.2	0.2	

Error Count Estimates for agegroup						
	1	2	3	4	5	Total
Rate	0.5062	0.4800	0.3736	0.8244	0.0000	0.4368
Priors	0.2000	0.2000	0.2000	0.2000	0.2000	

Conclusion

This project provided a comprehensive statistical evaluation of how mixture proportions and curing age influence concrete compressive strength. By applying descriptive analysis, correlation study, hierarchical clustering, regression modeling, logistic classification, and discriminant analysis, we gained a multi-dimensional understanding of the structural behavior of concrete mixes. Across all analyses, the results were consistent with engineering expectations and highlighted the central role of mix design in determining performance.

Key Takeaways

- **Strength Distribution & Behavior**
 - Strength spans a wide range from **2.3 MPa to 82.6 MPa**, demonstrating substantial variability across mixes.
 - The distribution is **right-skewed**, with most samples falling between **20–50 MPa** and a smaller group achieving high-performance strengths above **70 MPa**.
 - Early-age concrete is significantly weaker, with strength rising sharply in the first two months before stabilizing after ~90 days.
- **Influence of Mix Proportions**
 - **Cement/water ratio** shows the strongest positive relationship with strength ($r \approx 0.56$).
 - **Superplasticizer** and **slag** also contribute positively to strength development.
 - **Fly ash** and **fine aggregate** show mild negative associations, slightly reducing strength at higher proportions.
 - Composition ultimately plays a **larger role than curing age** once concrete reaches maturity.
- **Cluster Structure**
 - Ward's clustering identified **four distinct groups** representing different mixture strategies.
 - ANOVA confirmed **significant differences** across clusters ($p < 0.0001$).
 - Cluster 1 formed a clearly **higher-strength group**, while Clusters 2, 3, and 4 showed comparable strength levels.
- **Regression Insights (Age ≥ 100 days)**
 - Stepwise regression selected cement, slag, and superplasticizer as the most influential predictors.
 - The model achieved an **R^2 of ~ 0.56** , indicating moderate explanatory power.
 - Age was not selected, validating that long-term strength differences are almost entirely driven by composition.
- **Logistic Prediction of 50 MPa Strength**
 - The logistic model achieved **$\sim 95\%$ accuracy** and an **AUC of $\sim 0.90\text{--}0.93$** for predicting whether concrete aged 90–100 days reaches 50 MPa.
 - Cement, slag, and superplasticizer significantly increased the likelihood of reaching the structural threshold.

- **Age Group Classification (LDA)**

- LDA achieved **65–70% overall classification accuracy**.
- Early-age (<28 days) and long-cured (≥ 180 days) concrete were best classified, while mid-age groups overlapped.
- Strength and ratios together capture curing progress but cannot perfectly distinguish intermediate stages.

Overall Interpretation

In summary, this project shows that **mix design is the primary driver of concrete performance**, especially after curing stabilizes. Increasing cementitious content, balancing SCMs, and using admixtures appropriately leads to consistently higher strengths, while aggregate proportions must be managed carefully. Curing age remains critical in the early phase but less influential after maturity. The combination of clustering, regression, and classification techniques provided a robust, data-driven view of how mixture ratios govern strength, supporting practical decision-making in concrete design.
