Reaction Time Across a Training Programme

N = 90 · Dataset: repeated_training.csv

Statistical Tests Used

Objective	Test	Variables	Why this test
To describe the reaction_time_ms of the participants.	Descriptive Statistics	reaction_time_ms	Summary statistics for the continuous variables.
To determine whether reaction_time_ms differs across the repeated sessions (baseline, week 4, and week 8).	Repeated-Measures ANOVA	reaction_time_ms, session	Comparing 'reaction_time_ms' across 3 related conditions of 'session'.

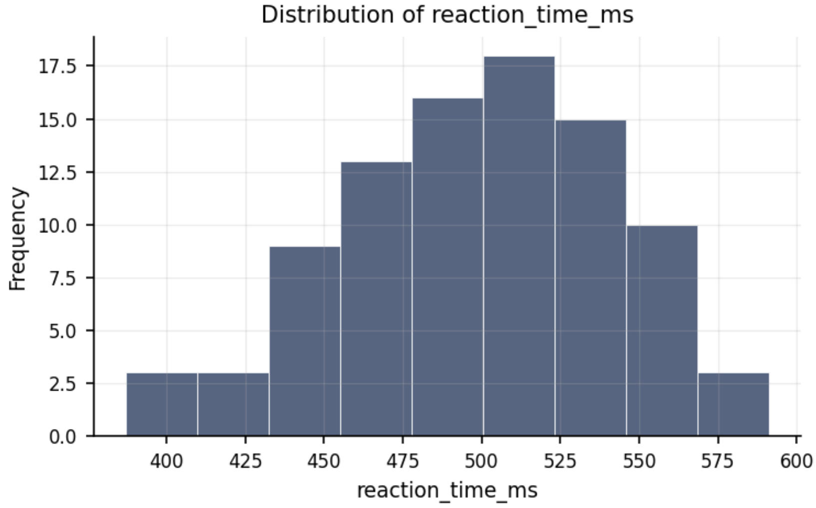
Methods

All statistical analyses were performed in Python using the SciPy and statsmodels libraries via thericerca's deterministic analysis engine (v0.1.0). Statistical significance was evaluated at $\alpha = .05$ (a 95% confidence level). Descriptive statistics were computed to summarise the study's variables. To assess differences in a continuous outcome measured on the same subjects across three or more conditions, a one-way repeated-measures analysis of variance (ANOVA), with the Greenhouse–Geisser correction when sphericity was violated was used (Greenhouse & Geisser, 1959). For each inferential test, the appropriate effect size is reported with its 95% confidence interval, alongside the achieved statistical power to detect the observed effect (Cohen, 1988).

Results

To describe the reaction_time_ms of the participants.

This simply summarises the data — the average, spread, and range of each variable are in the table above.



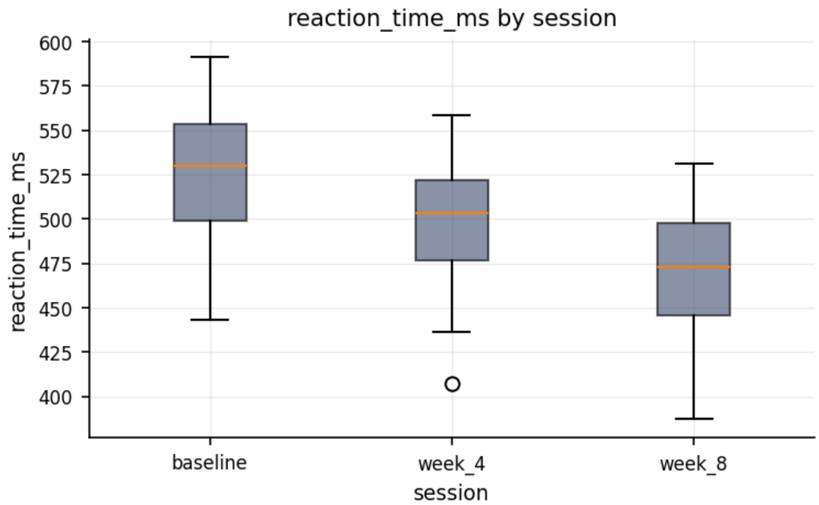
Distribution of reaction_time_ms.

max	mean	median	min	mode	n	range	std	variable	variance
591.2	498.759	501.45	387.1	446.2	90	204.1	42.785	reaction_time_ms	1830.57

To determine whether reaction_time_ms differs across the repeated sessions (baseline, week 4, and week 8).

This test found a statistically reliable pattern. The strength of this pattern is large.

Statistically: $F(2, 58) = 313.71$, $p < .001$, $\eta p^2 = .92$, achieved power = 1.00



Distribution of reaction_time_ms across session groups.

condition	mean	n	std
baseline	525.68	30	37.11
week_4	499.127	30	37.295
week_8	471.47	30	36.465

⚠ Sphericity was violated (Mauchly's test was significant); the Greenhouse-Geisser correction was applied to the degrees of freedom, and the corrected p-value is reported alongside the uncorrected result.

Discussion

Repeated-Measures ANOVA. The Repeated-Measures ANOVA produced a statistically reliable result in this sample.

This study focused on evaluating how participants' reaction times, measured in milliseconds, shifted across three specific testing sessions: an initial baseline assessment, a mid-point evaluation at week four, and a final assessment at week eight. The core objective was to determine whether reaction speeds remained static or experienced meaningful changes over time. The analysis demonstrated a clear and pronounced pattern of change across these sessions. In practical terms, calling a finding statistically significant means that the observed differences in reaction time across the baseline, week four, and week eight sessions are extremely unlikely to be the result of random chance or everyday variability. Instead, the data reflects a genuine, systematic shift in performance across the study period. Moreover, the magnitude of this effect—a measure showing how strongly the variables are connected—was remarkably large. This indicates that the progression across testing weeks accounts for a major share of the total differences seen in reaction times, demonstrating that time and repeated assessment play a critical role in reaction speed outcomes.

Conclusions

In conclusion, the study successfully fulfilled its primary purpose by showing that reaction times change in a strong and highly consistent manner across baseline, week four, and week eight sessions. Because these findings are statistically reliable, readers can be confident that the observed performance trends reflect a true underlying pattern rather than a random artifact of the sample. The substantial proportion of variance explained by the repeated sessions reinforces that these shifts are not merely noticeable, but practically meaningful. Altogether, the results offer a coherent picture of how reaction times evolve across structured time intervals, providing a solid factual foundation to guide decision-making and future study design, while remaining contextualised by the specific participants and conditions involved.

Recommendations

Based on these outcomes, several practical steps are recommended. First, because the changes in reaction time across the three sessions are so robust, future initiatives should replicate this protocol in independent samples to confirm that these patterns hold true across different groups and settings. Additionally, researchers interested in the underlying drivers of these changes should utilize longitudinal or experimental frameworks designed to determine whether improvements stem from learning effects, physical adaptation, or specific external interventions. Throughout future research, analysts should continue reporting effect magnitudes alongside basic significance checks to ensure that the practical importance of any shift remains clear. Finally, maintaining strict, consistent testing protocols across all time points and screening data thoroughly for missing records or extreme values will ensure that ongoing monitoring remains reliable and actionable.

Limitations

A few important limitations should be acknowledged alongside these conclusions. First, while data quality procedures detected no major flaws or data anomalies, the study design observes trends over time rather than manipulating every potential influence. As a result, showing a clear statistical link between the passage of time and faster or slower reaction times does not automatically rule out unmeasured influences, such as individual motivation, external lifestyle factors, or natural familiarity with the testing tool. Second, these findings directly describe the specific group of individuals who participated in the study. Extending these conclusions to a broader population assumes that this sample accurately represents the wider group, meaning caution should be exercised when applying these exact expectations to different demographics or settings.

References

1. Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. Psychometrika, 24(2), 95–112.
2. Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Lawrence Erlbaum Associates.
3. Seabold, S., & Perktold, J. (2010). statsmodels: Econometric and statistical modeling with Python. Proceedings of the 9th Python in Science Conference, 92–96.
4. Virtanen, P., Gommers, R., Oliphant, T. E., et al. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. Nature Methods, 17, 261–272.

Python Code Appendix

Descriptive Statistics

```
df[[]].describe()
```

Numeric-claim validation: PASSED (13 numbers checked). Engine v0.1.0 · Registry v0.2.0. Prose source: llm:gemini.