

Vollständige Bewertung der Mathematics-AA-SL-Exploration

„Formulating a Suitable Equation to Model the Cooling Process of Hibiscus
Tea“

Strenge Moderationsperspektive | Ziel: bestmögliche Voraussetzungen für eine 7

Bewertungsstand: 27. August 2026

Geprüfte Fassung: 22 PDF-Seiten, davon 2 Seiten Literaturverzeichnis

Kurs: IB Mathematics: Analysis and Approaches, Standard Level

1. Gesamturteil

Die Exploration hat eine gute Grundidee, einen glaubwürdigen persönlichen Ausgangspunkt, selbst erhobene Daten und eine sinnvolle Verbindung von Exponentialfunktionen, Logarithmen, Regression, Differentialgleichungen und Modellkritik. Das ist eine deutlich bessere Ausgangslage als bei vielen durchschnittlichen IAs.

In der vorliegenden Fassung verhindern jedoch mehrere **entscheidende mathematische Inkonsistenzen** eine sichere Spitzennote. Besonders schwer wiegen:

1. In Gleichung C steht einmal $e^{-0.57955t}$ statt $e^{-0.057955t}$ - ein Faktor-10-Fehler.
2. Die Parameter von Gleichung B stimmen zwischen Regressionsgrafik, Text, Gleichung, Wertetabelle und RMSE nicht überein.
3. Die Modelle werden nicht auf derselben Datenmenge verglichen. Der Schluss, Gleichung C sei für den relevanten Bereich am besten, ist mit den angegebenen Daten falsch.
4. Das berichtete angepasste RMSE von $1.991 \square^{\circ} C$ ist nicht reproduzierbar. Aus Tabelle 1.7 ergibt sich für $t \geq 10$ ungefähr $2.050 \square^{\circ} C$.
5. „Boyle's gas law“ erklärt die Annäherung an die Raumtemperatur nicht.
6. Einheiten, Variablennamen, Achsenbeschriftungen und Genauigkeit sind mehrfach inkonsistent.

Mein strenges, aber realistisches Moderationsurteil lautet deshalb derzeit **15/20 Punkten**; je nach Gewichtung der Kommunikations- und Reflexionsfehler ist eine Spanne von **14-16/20** plausibel. Nach einer sauberen mathematischen Überarbeitung ist **18-20/20** erreichbar.

Wichtig: Eine Math-IA erhält keine eigenständige Fachnote von 1 bis 7. Sie wird mit maximal 20 Rohpunkten bewertet und zählt in Mathematics AA SL 20 % zur Fachnote. Ob die Gesamtnote 7 erreicht wird, hängt zusätzlich von den Prüfungen und den jeweiligen Session-Grenzen ab.

2. Verwendete aktuelle IB-Grundlage

Für eine Einreichung vor der ersten Prüfungsrunde Mai 2029 gelten weiterhin die fünf Kriterien des gegenwärtigen Mathematics-AA-Kurses:

Kriterium	Maximum	Kernanforderung
A: Presentation	4	kohärent, gut organisiert und prägnant
B: Mathematical communication	4	relevante, angemessene und durchgehend konsistente mathematische Kommunikation

Kriterium	Maximum	Kernanforderung
C: Personal engagement	3	eigenständige, authentische und die Untersuchung vorantreibende Auseinandersetzung
D: Reflection	3	substanzielle kritische Reflexion im Verlauf der Exploration
E: Use of mathematics - SL	6	relevante, korrekte Mathematik auf Kursniveau mit gründlichem Verständnis

Die neue Kriterienstruktur mit „Problem specification, Abstraction, Computation, Interpretation“ gilt erst für den neuen Kurs mit **erstem Unterricht ab August 2027 und erster Prüfung im Mai 2029**. Sie darf für eine gegenwärtige IA nicht anstelle der oben genannten Kriterien verwendet werden.

3. Voraussichtliche Bewertung nach allen Kriterien

Kriterium	Prognose	Begründung	Weg zur Höchstpunktzahl
A: Presentation	3/4	Klarer Aufbau von Fragestellung bis Schluss; Grafiken stehen meist an der richtigen Stelle. Nicht vollständig prägnant: drei sehr ähnliche Parameterschätzungen werden lang dargestellt, Nummerierungen sind inkonsistent, Seitenzahl und Seitennavigation sind fehlerhaft.	Wiederholungen kürzen, Tabellen zusammenführen, Seitenzahlen ergänzen, Abbildungs- und Tabellenverweise vereinheitlichen, zentrale Modellentscheidung früher und klarer führen.
B: Mathematical communication	2/4	Viele passende Darstellungsformen, aber mehrere wesentliche Inkonsistenzen: $t = 22.4$ statt $T = 22.4$, Sekunden statt Minuten, Faktor-10-Fehler, fehlende Parametereinheiten, Symbolkonflikt m , unterschiedliche Regressionsparameter.	Eine eindeutige Symboltabelle verwenden; jede Gleichung, Achse, Tabelle und Einheit prüfen; alle Zahlen aus einer einzigen reproduzierbaren Regressionsrechnung ableiten.
C: Personal engagement	2/3	Persönliche Teeroutine, eigener Schwellenwert und eigene Messungen sind authentisch. Ein Teil der Entwicklung wirkt jedoch nachträglich „für die IA gebaut“, besonders der inszenierte Übergang von Versuch A über B zu C und die Auswahl zweier günstiger Punkte.	Eine echte Entscheidung sichtbar machen: Vorhersage formulieren, Residuen untersuchen, eine Parameterschätzung begründet verwerfen, mindestens einen Trial als Validierung verwenden.
D: Reflection	3/3	Es gibt wiederholt gute Reflexion:	Die erkannte Schwäche tatsächlich

Kriterium	Prognose	Begründung	Weg zur Höchstpunktzahl
	knapp	Transformationsfehlergewichtung, unmögliche Anfangstemperatur, Modellgrenzen, ungleicher Vergleich und mögliche Verbesserungen werden angesprochen. Die Qualität wird allerdings dadurch geschwächt, dass der erkannte unfaire Vergleich trotzdem als Entscheidungsgrundlage dient.	mathematisch beheben; Konsequenzen jeder Einschränkung quantifizieren statt nur aufzuzählen.
E: Use of mathematics - SL	5/6	Relevante und überwiegend richtige Mathematik; die Differentialgleichung liegt sogar über dem typischen SL-Niveau und wird größtenteils verstanden. Der falsche Exponent, die inkonsistente Regression, das falsche RMSE und die nicht haltbare Modellwahl verhindern „durchgehend korrekt“.	Nichtlineare Regression direkt in $\square^{\circ}C$, fairer Vergleich, reproduzierbare Residuen/RMSE, Unsicherheits- oder Sensitivitätsanalyse und fehlerfreie Notation.
Gesamt	15/20	Gute, aber noch nicht sichere Top-IA.	Realistisches Ziel nach Überarbeitung: 19/20.

4. Vollständige, durchnummerierte Fehler- und Risikotabelle

4.1 Form, Aufbau, Datenerhebung und Darstellung

Nr.	Fundstelle	Fehler bzw. Risiko und warum es relevant ist	Konkrete Korrektur oder Alternativtext
1	Titelseite	„15 pages“ stimmt nicht mit der Datei überein. Die PDF hat 22 Seiten; ohne das zweiseitige Literaturverzeichnis sind es 20 Seiten. Eine falsche Seitenangabe wirkt unkontrolliert und betrifft Kriterium A/B.	Seitenzahl nach der Endfassung neu zählen. Zum Beispiel: “20 pages excluding the bibliography” .
2	Gesamtdokument	Es fehlen sichtbare Seitenzahlen. Bei einer langen Exploration erschwert das Moderation, Querverweise und Feedback.	Seitenzahlen in die Fußzeile aufnehmen; Titelseite optional ohne sichtbare Zahl, aber mitzählen.
3	Titelseite	„An exploration into Cooling Process“ ist grammatisch unvollständig und unnatürlich.	“An exploration of the cooling process of hibiscus tea” oder die Zeile ganz entfernen.
4	Gesamtdokument	Die Arbeit liegt mit 20 Seiten ohne Bibliografie an der oberen empfohlenen Grenze und wirkt teilweise einzeilig. Entscheidend ist zwar Qualität statt Länge, doch unnötige Wiederholung gefährdet die 4 in Kriterium A.	Rohdatentabellen zusammenführen, elementare Transformationen kürzen und drei nahezu gleiche Modellgrafiken auf eine Vergleichsgrafik reduzieren. Ziel: etwa 15-17 gehaltvolle Seiten ohne Inhaltsverlust.

Nr.	Fundstelle	Fehler bzw. Risiko und warum es relevant ist	Konkrete Korrektur oder Alternativtext
5	Tabellen 1-1.9	Die Nummerierung mischt „Table 1“, „Table 1.1“, „Table 1.2“ usw. Das wirkt wie Untertabellen, obwohl es eigenständige Tabellen sind.	Fortlaufend Table 1, Table 2, ... verwenden oder bewusst Table 1a-1d für zusammengehörige Datentabellen.
6	Figures 1-1.7	Im Text stehen „Figure 1.0“ und „Figure 1“, während die Beschriftungen „Figure 1“, „Figure 1.1“ usw. verwenden.	Eine fortlaufende Nummerierung wählen und jeden Verweis exakt daran anpassen.
7	Forschungsfrage	„for say 300 mL“ enthält Füllsprache und klingt nicht wissenschaftlich präzise.	“How accurately does Newton’s law of cooling model the cooling of 300 mL of hibiscus tea, and what time does the model predict for the temperature to fall below 45° C in a room at 22.4° C?”
8	Forschungsfrage	„How accurately“ ist nicht operationalisiert. Es bleibt offen, ob Genauigkeit über RMSE, Schwellenzeit, Residuen oder einen Validierungs-Trial bewertet wird.	Direkt nach der Frage festlegen: “Accuracy will be assessed using RMSE in degrees Celsius, residual plots, and the error in the predicted 45° C crossing time on a separate validation trial.”
9	Introduction	„Temperature is continuous“ wird als Tatsache statt als Modellannahme präsentiert; gemessen wird diskret.	“I modelled temperature as a continuous function of time, although the experimental readings were taken at discrete five-minute intervals.”
10	Introduction	„the surrounding room places a lower restriction“ ist physikalisch zu absolut; Verdunstung kann eine Flüssigkeit unter die trockene Umgebungstemperatur abkühlen.	“Under the simplified Newton cooling model, and neglecting evaporative cooling below ambient temperature, the room temperature is the long-term equilibrium value.”
11	Introduction	Grammatik: „The bright red colour and sour taste makes“. Ein zusammengesetztes Subjekt verlangt „make“.	“The bright red colour and sour taste make the drink especially enjoyable.”
12	Introduction	„thermal equilibrium“ ist falsch geschrieben.	“thermal equilibrium”.
13	Introduction	Die Begründung „hibiscus tea is less viscous than coffee“ ist unbelegt und stellt keinen nachvollziehbaren Zusammenhang zur bevorzugten Trinktemperatur her.	Entfernen oder ehrlich persönlich formulieren: “I chose 45° C from repeated personal trials; it is therefore a subjective usability threshold rather than a universal preference.”
14	Background	Der Absatz zur Anthocyanin-Extraktion trägt kaum zur mathematischen Fragestellung bei und kostet Prägnanz.	Auf einen Satz kürzen: “I used the same hot-infusion procedure in each trial so that preparation differences would not dominate the cooling data.”
15	Methodology	Die drei tatsächlichen Raumtemperaturen werden nicht als Rohdaten gezeigt. Damit ist der Mittelwert 22.4° C nicht reproduzierbar.	Für jeden Trial Anfangs-, End- und Mittelwert der Raumtemperatur angeben oder die Zeitreihe in einer kleinen Tabelle zeigen.
16	Methodology	„A fluctuation of about $\pm 0.3^{\circ} \text{C}$ “ ist mehrdeutig: Messunsicherheit, beobachtete Spannweite oder Abweichung vom Mittelwert?	“During the trials the room reading ranged from ... to ... $\square^{\circ} \text{C}$; the thermometer resolution was ...

Nr.	Fundstelle	Fehler bzw. Risiko und warum es relevant ist	Konkrete Korrektur oder Alternativtext
17	Methodology	„Using three trials should eliminate random errors“ ist falsch. Wiederholungen beseitigen Zufallsfehler nicht, sondern ermöglichen eine Schätzung und können ihren Einfluss auf den Mittelwert reduzieren.	“Three repeated trials do not remove random error, but they allow the variability of the measurements to be estimated and reduce its effect on the mean.”
18	Methodology	Auflösung und Unsicherheit werden vermischt. Eine Anzeige auf $0,1^{\circ}\text{C}$ ist nicht automatisch eine Messunsicherheit von $\pm 0,1^{\circ}\text{C}$.	Herstellerangabe oder Kalibrierung nennen; sonst vorsichtig zwischen resolution und estimated measurement uncertainty unterscheiden.
19	Methodology	Zeitmessungsunsicherheit, Verzögerung beim Ablesen und mögliche Probe-Trägheit fehlen. Diese Effekte sind gerade nahe der Schwelle relevant.	Timerauflösung, typische Ablesedauer und Ansprechzeit der Sonde nennen; Einfluss qualitativ oder per Sensitivitätsanalyse prüfen.
20	Raw data	Drei nahezu identische Tabellen verbrauchen viel Raum und erschweren den direkten Vergleich.	Eine Tabelle mit Spalten T_1, T_2, T_3 , Mittelwert und Standardabweichung erstellen.
21	Processed data	Es wird nur der Mittelwert berichtet. Ohne Standardabweichung oder Fehlerbalken ist die Reproduzierbarkeit der Trials unsichtbar.	Für jede Zeit S_T berechnen und in Tabelle/Grafik als Fehlerbalken zeigen.
22	Processed data	Erst zu mitteln und danach nur die Mittelkurve zu modellieren versteckt die Variabilität der Parameterschätzungen und der Schwellenzeit.	Zusätzlich jedes Trial separat fitten oder mindestens die 45°C -Zeit pro Trial berechnen.
23	Figure 1	Nur Farbe trennt Punkte und Asymptote. Das ist für Graustufendruck und Farbsehschwäche ungünstig.	Unterschiedliche Marker und Linienarten verwenden; Legende nicht nur über „red/black“ beschreiben.
24	Rohdaten/ Methodik	Luftfeuchte wird als nicht kontrollierbar genannt, aber nicht gemessen. Da Verdunstung als Erklärung verwendet wird, bleibt diese Behauptung unbelegt.	Entweder Luftfeuchte messen oder Verdunstung nur als mögliche, nicht nachgewiesene Ursache bezeichnen.

4.2 Mathematik, Modellierung und Statistik

Nr.	Fundstelle	Fehler bzw. Risiko und warum es relevant ist	Konkrete Korrektur oder Alternativtext
25	Nach Figure 1	„A horizontal asymptote at $t = 22.4$ “ verwechselt Zeit und Temperatur.	“A horizontal asymptote at $T = 22.4^{\circ}\text{C}$ was added.”
26	Nach Figure 1	Boyle's law erklärt diesen Sachverhalt nicht. Boyle beschreibt bei konstanter Temperatur die Beziehung zwischen Druck und Volumen eines Gases.	“In the simplified Newton cooling model, the net heat transfer tends to zero as T approaches the ambient temperature T_a .”
27	Nach Figure 1	„a substance cannot decrease its temperature below that of its environment“ ist zu absolut und widerspricht möglichen Verdunstungseffekten.	“Without an additional cooling mechanism such as evaporation, the Newton model approaches T_a asymptotically and does not cross it.”

Nr.	Fundstelle	Fehler bzw. Risiko und warum es relevant ist	Konkrete Korrektur oder Alternativtext
28	Nach Figure 1	„negative correlation“ ist für eine gekrümmte Zeitreihe unpräzise; ein Korrelationskoeffizient wird nicht berechnet.	“The data show a monotonic, nonlinear decrease in temperature with time.”
29	Attempt A	Der Einstieg $T(t) = e^{-t}$ ist dimensionsmäßig unvollständig: Die rechte Seite ist dimensionslos, T hat Temperatureinheiten.	Als reine Form zuerst $f(t) = e^{-t}$ schreiben; erst nach Skalierung $T(t) = T_a + A e^{-kt}$ verwenden, wobei A in $^{\circ}\text{C}$ und k in min^{-1} angegeben werden.
30	Attempt A	k wird als „horizontal stretch factor“ bezeichnet. In e^{-kt} ist $1/k$ die charakteristische Zeitskala; k ist eine Rate.	“The cooling-rate constant k has units min^{-1}; the characteristic time is $\tau = 1/k$.”
31	Attempt A	m wird als vertikaler Streckfaktor behandelt, aber seine Einheit fehlt.	Besser A verwenden und definieren: “$A = T(0) - T_a$ is the initial excess temperature, measured in $^{\circ}\text{C}$.”
32	Attempt A	„first and last observations ... should create the most representative model“ ist nicht begründet. Endpunkte können Messfehler enthalten; Zwischenwerte werden ignoriert.	“This endpoint estimate captures the total observed change, but it is sensitive to both endpoint measurements and is used only as a preliminary benchmark.”
33	Determining constants A	„utmost accuracy“ und sechs Dezimalstellen sind bei auf 0.1°C gerundeten Temperaturen eine Scheingenauigkeit.	Intern ungerundet rechnen, Ergebnisse aber angemessen berichten, z. B. $k = 0.0473 \text{ min}^{-1}$ und die verwendete Rechengenauigkeit erklären.
34	Logarithm steps	$\ln(T - T_a)$ nimmt formal den Logarithmus einer dimensionsbehafteten Größe. Die Aussage „logarithms have no units“ löst dieses Problem nicht.	Eine dimensionslose Größe definieren: $\theta = (T - T_a)/(1^{\circ}\text{C})$ und dann $\ln \theta$ verwenden. Für SL genügt alternativ die klare Erklärung, dass die numerischen Werte der Temperaturdifferenz in $^{\circ}\text{C}$ logarithmiert werden.
35	Attempt B	Das Symbol m steht bereits für die Exponentialamplitude und wird gleichzeitig in $y = mx + c$ als Steigung benutzt. Das macht die Zuordnung unnötig mehrdeutig.	Für die Gerade $y = bt + c$ verwenden: $b = -k$ und $c = \ln A$.
36	Attempt B	Die Zeile „ $mx = kt$ “ hat das falsche Vorzeichen und ist keine saubere Variablenzuordnung.	Schreiben: $y = \ln(T - T_a)$, $x = t$, $b = -k$, $c = \ln A$, also $y = bx + c$.
37	Table 1.5	Der Tabellenkopf lautet „Time (s)“, obwohl alle Messungen und Rechnungen Minuten verwenden.	“Time (min)”.
38	Attempt B	„If T_a were changed for every row, ... would no	“Using a time-dependent ambient temperature would lead to a non-autonomous differential

Nr.	Fundstelle	Fehler bzw. Risiko und warum es relevant ist	Konkrete Korrektur oder Alternativtext
		longer describe one function“ ist mathematisch falsch. Ein zeitabhängiges $T_a(t)$ ergibt weiterhin eine Funktion, nur nicht mehr diese einfache linearisierte Form.	equation and would invalidate this simple linearisation.”
39	Figure 1.3	„98.71% of the variation in temperature is explained by time“ interpretiert R^2 falsch. Die Regression bezieht sich auf $\ln(T - T_a)$, nicht auf T .	“ $R^2 = 0.9870$ means that 98.70% of the variation in the transformed response $\ln(T - T_a)$ is accounted for by its linear relation with time in this data set.”
40	Figure 1.3	Grammatik: „a exponential equation“.	“an exponential equation”.
41	Regression B	Die Grafik zeigt ungefähr $y = -0.05100x + 4.20476$, der Text verwendet 4.20430 und $A = 66.9735$. Aus den in Tabelle 1.3 gerundeten Werten folgt dagegen $A = e^{4.204761} \approx 67.0046$.	Eine einzige ungerundete Regressionsausgabe verwenden. Bei den sichtbaren Mittelwerten: $T(t) = 22.4 + 67.0046 e^{-0.0509973t}$.
42	Equation B/Table 1.7	Die Werte in Tabelle 1.7 entsprechen der Grafik-Version von B mit $A \approx 67.0046$, nicht der gedruckten Equation B mit $A = 66.9735$. Damit sind Gleichung, Tabelle und RMSE nicht gegenseitig reproduzierbar.	Tabelle und RMSE automatisch aus derselben finalen Zellformel bzw. Gleichung erzeugen; Parameter nicht manuell an mehreren Stellen abtippen.
43	Tabellen 1.7/Parameter	Vier bis sechs Nachkommastellen bei Vorhersagetemperaturen suggerieren eine Genauigkeit, die Messgerät und Modell nicht besitzen.	Für Rechenkontrolle intern mehr Stellen behalten, im Fließtext etwa 0.1°C oder 0.01°C zeigen und Unsicherheit ausdrücklich nennen.
44	Nach Figure 1.2	Ein „repeated pattern of residuals“ wird behauptet, aber kein Residuenplot gezeigt. Kriterium E/D verlangt hier Evidenz.	Für jedes Modell einen Plot $r_i = T_i - \hat{T}_i$ gegen t_i ergänzen, Nulllinie einzeichnen und Muster interpretieren.
45	Nach Figure 1.2	Luftbewegung, thermische Masse und Verdunstung werden als Ursachen der Abweichung genannt, ohne dass sie gemessen oder getrennt wurden.	“Possible explanations include ..., but the present data cannot identify which mechanism caused the residual pattern.”
46	Newton's law	Die Einheit von k fehlt. Dadurch ist auch die Dimensionsprüfung der Differentialgleichung unvollständig.	Direkt definieren: $k > 0$ in min^{-1} , T und T_a in $^\circ\text{C}$, t in Minuten.
47	Differentialgleichung	Aus $\ln(T - T_a) = -kt + C$ wird ohne Begründung $T - T_a = e^C e^{-kt}$. Das Vorzeichen aus dem Betrag wird stillschweigend verworfen.	“Because all measured temperatures satisfy $T > T_a$, $T - T_a$ is positive; therefore the absolute value may be removed on the observed domain.”
48	Lösung der DGL	Der freie Parameter Z wird eingeführt, obwohl seine Bedeutung durch die Anfangsbedingung	$A = T(0) - T_a$ verwenden. Dann lautet die

Nr.	Fundstelle	Fehler bzw. Risiko und warum es relevant ist	Konkrete Korrektur oder Alternativtext
		direkt klar ist. Ohne Einheit bleibt er abstrakt.	Lösung $T(t) = T_a + (T_0 - T_a)e^{-kt}$.
49	Attempt C	Die Begründung der Punkte $t = 15$ und $t = 45$ ist sachlich schwach. $T(45) = 28.2^\circ \text{C}$ liegt nicht „close to the drinking threshold“ von 45°C .	Wenn eine schwelennahe Zweipunktschätzung gewünscht ist, die Punkte bei 20 und 25 Minuten verwenden - besser jedoch alle Daten per nichtlinearer Regression fitten.
50	Attempt C	Die zwei Punkte werden nach Sichtung der Daten ausgewählt. Das erzeugt Auswahlbias und lässt C im gewählten Bereich künstlich gut aussehen.	Auswahlregel vor der Auswertung festlegen oder einen Trainings-/Validierungsansatz verwenden.
51	Equation C	Schwerer Tipp-/Rechenfehler: Gedruckt ist $T(t) = 78.7149e^{-0.57955t} + 22.4$. Verwendet wird tatsächlich $-0.057955t$. Der gedruckte Exponent ist zehnfach zu groß.	Überall einheitlich: $T_C(t) = 22.4 + 78.7149e^{-0.057955t}$. Danach sämtliche Tabellen und Grafiken neu generieren.
52	Figure 1.5	Die x-Achse ist „Time (seconds)“, obwohl die Zahlen Minuten sind. Der erläuternde Satz behebt die falsche Grafik nicht.	Achse in der Grafik selbst auf Time (minutes) korrigieren.
53	Modelle A-C	A, B und C sind nicht drei verschiedene physikalische Modelle. Alle haben die Form $T = T_a + Ae^{-kt}$; nur die Parameterschätzung ist verschieden.	Überschrift ändern zu “Comparing three parameter-estimation methods for the same Newton cooling model” .
54	Modellierung	Gerade weil RMSE in $^\circ \text{C}$ das Vergleichsmaß ist, fehlt die naheliegende Methode: direkte nichtlineare kleinste Quadrate in der ursprünglichen Temperaturskala.	$SSE(A, k) = \sum_i \left[T_i - (T_a + Ae^{-kt_i}) \right]^2$ minimieren und das Ergebnis mit der Log-Regression vergleichen.
55	RMSE section	RMSE ist nicht allgemein „the standard deviation of the residuals“. Das gilt nur unter zusätzlichen Bedingungen und mit anderer Zentrierung/Normierung.	“RMSE is the square root of the mean squared residual and measures the typical prediction error in the original temperature units.”
56	RMSE section	„a lower result shows that the regression line is a better suit“ ist sprachlich und sachlich ungenau; nicht alle verglichenen Objekte sind Regressionsgeraden.	“When models are evaluated on the same observations, a lower RMSE indicates a closer fit in degrees Celsius.”
57	RMSE calculations	Es fehlt eine beispielhafte RMSE-Rechnung oder eine eindeutige Spreadsheet-Formel. Dadurch fällt der falsche angepasste Wert nicht auf.	Eine Residuentabelle oder mindestens eine vollständig eingesetzte Rechnung zeigen; Dateiformel/Software nennen.
58	Table 1.9	Der Wert 1.991°C für C bei $t \geq 10$ ist aus den angegebenen Werten nicht reproduzierbar. Mit Tabelle 1.7 ergibt sich ungefähr 2.050°C .	RMSE neu berechnen: $\sqrt{\frac{1}{11} \sum_{t=10}^{60} (T_i - \hat{T}_i)^2} \approx 2.050^\circ \text{C}.$

Nr.	Fundstelle	Fehler bzw. Risiko und warum es relevant ist	Konkrete Korrektur oder Alternativtext
59	Adapted comparison	C auf $t \geq 10$ wird mit A und B auf $t \geq 0$ verglichen. RMSE-Werte auf unterschiedlichen Beobachtungsmengen sind nicht direkt vergleichbar.	Alle Gleichungen auf demselben Bereich prüfen. Für $t \geq 10$ ergeben sich aus den sichtbaren Daten ungefähr: A 3.243, B 1.604, C 2.050 $^{\circ}\text{C}$.
60	Auswahl C	Nach dem fairen Vergleich ist B, nicht C, unter den drei vorhandenen Gleichungen im Bereich $t \geq 10$ am besten. Die zentrale Modellentscheidung der IA ist daher nicht durch die eigenen Daten gedeckt.	Entweder B auswählen oder - besser - direkte nichtlineare Regression einsetzen und auf einer separaten Messreihe validieren.
61	Interpretation C	Die hohe Anfangstemperatur von C entsteht vor allem durch die Zweipunktschätzung, nicht primär dadurch, dass $T_a = 22.4^{\circ}\text{C}$ konstant gesetzt wurde.	“The unrealistic intercept is a consequence of fitting only the 15- and 45-minute observations and extrapolating the resulting curve back to $t = 0$.”
62	Interpretation	Die Erwärmung der lokalen Luft und die Wärmerückgabe der Tasse werden als feststehende Erklärungen formuliert. Die Daten trennen diese Mechanismen nicht.	Als Hypothesen kennzeichnen und ggf. ein gekoppeltes Tasse-Flüssigkeit-Modell nur als Erweiterung diskutieren.
63	Ausschluss 0-10 min	„I would never drink the tea immediately“ rechtfertigt nicht automatisch das Entfernen der frühen Werte, denn die gesuchte Zeit wird ab $t = 0$ gemessen.	Relevanten Auswertebereich vorab definieren und trotzdem den Zeitsprung beibehalten; alle Methoden auf demselben Bereich fitten und testen.
64	Validierung	Alle Daten werden zugleich zur Wahl/Fit des Modells und zur Bewertung benutzt. Das ist In-Sample-Güte, keine unabhängige Vorhersagegenauigkeit.	Trial 1 und 2 zum Fitten, Trial 3 als Validierung verwenden; anschließend Rollen rotieren oder einen vierten Validierungs-Trial erheben.

4.3 Schlussfolgerung, Reflexion, Quellen und Sprache

Nr.	Fundstelle	Fehler bzw. Risiko und warum es relevant ist	Konkrete Korrektur oder Alternativtext
65	Answering the question	Der Modellwert 21.53 min und die lineare Interpolation 24.00 min unterscheiden sich um 2.47 min. „reasonably close“ bleibt ohne Quantifizierung zu vage.	“The model predicts 21.53 min, whereas interpolation of the adjacent observations gives 24.00 min; the model predicts the crossing 2.47 min earlier.”
66	Schluss	„maximum amount of time“ klingt exakter, als Daten und Modell es erlauben.	“The selected model gives an estimated crossing time of about 22 minutes; the observed five-minute data place the crossing between 20 and 25 minutes.”
67	Schluss	„If I begin drinking within twenty minutes, the model predicts the tea will remain above ...“ ist mehrdeutig: oberhalb der Schwelle beim Beginn oder während des gesamten Trinkens?	“The model predicts that the tea is still above 45°C at $t = 20$ min.”
68	Ergebnis	Es gibt kein Unsicherheitsintervall für die Schwellenzeit. Eine einzelne Dezimalzahl ist bei	Interpolation pro Trial berichten: ungefähr 24.15,

Nr.	Fundstelle	Fehler bzw. Risiko und warum es relevant ist	Konkrete Korrektur oder Alternativtext
		5-Minuten-Abständen und Modellbias nicht ausreichend.	23.75 und 24.10 min; Mittelwert 24.00 min, Stichproben-SD etwa 0.22 min. Zusätzlich Modellunsicherheit und systematischen Unterschied diskutieren.
69	Limitations	„More than three trials ... permit uncertainty intervals“ ist falsch. Schon drei Trials erlauben eine grobe Streuungsschätzung; mehr Trials machen sie zuverlässiger.	“Three trials allow a preliminary estimate of variability; additional trials would make an uncertainty interval more reliable.”
70	Limitations	„Recording every minute“ ist sinnvoll, aber nicht auf die Fragestellung zugeschnitten.	Besonders zwischen 18 und 28 Minuten im Abstand von 30-60 s messen; außerhalb dieses Bereichs können größere Intervalle genügen.
71	Limitations	„measure mass rather than volume because evaporation reduced the amount“ ist unvollständig. Nur die Anfangsmasse zu messen zeigt den Verdunstungsverlauf nicht.	Tasse samt Inhalt während des Versuchs wiegen oder Anfangs- und Endmasse vergleichen; Messung darf den Wärmefluss nicht wesentlich verändern.
72	Limitations	T_a als Funktion der Zeit vorzuschlagen ist mathematisch anspruchsvoll, aber die daraus entstehende Differentialgleichung wird nicht angegeben. Der Vorschlag bleibt dekorativ.	Entweder eine Sensitivitätsanalyse mit $T_a = 22.1, 22.4, 22.7^\circ \text{C}$ durchführen oder die Gleichung $\frac{dT}{dt} = -k(T - T_a(t))$ samt Lösungsstrategie erläutern.
73	Limitations	Ein Magnetrührer verbessert Homogenität, verändert aber zugleich die Konvektion und damit k .	Dies ausdrücklich als neues kontrolliertes Versuchsdesign behandeln, nicht als neutrale Korrektur des bestehenden Experiments.
74	Reflection	Der kurze Abschnitt „Reflection“ fasst vor allem Methoden zusammen. Die stärkste Reflexion steht verteilt in anderen Abschnitten.	Den Abschnitt nutzen, um eine echte Erkenntnis zu formulieren: Warum hat hohe linearisierte R^2 -Güte nicht zum kleinsten Temperatur-RMSE geführt, und wie ändert das die Modellwahl?
75	IA-Ton	„Since both equations were created mainly from the shape ... I completed additional research“ klingt wie eine nachträglich inszenierte Rubrik-Erzählung.	“The residual pattern suggested that a fitted exponential alone did not explain why this form should arise, so I derived the model from a rate equation.”
76	IA-Ton	„The mathematics developed from a visual conjecture into a physically motivated model“ klingt sehr nach Bewertungsraster und wenig nach eigener Denkbewegung.	“At first I chose the exponential form from the graph. Deriving Newton's law later showed which assumptions were hidden in that choice.”
77	IA-Ton	„This initially suggests ...“, „On closer observation ...“, „Nevertheless, it answers the investigation more directly“ wirken formelhaft, besonders weil die mathematische Entscheidung nicht folgt.	Direkte Aussagen mit Zahlen verwenden: “On $t \geq 10$, model B has RMSE 1.604°C and model C 2.050°C; therefore B fits this domain more closely.”
78	IA-Ton	„The investigation has transformed an ordinary routine into a practical use of ...“ ist eine	“For my preparation method, the data place the

Nr.	Fundstelle	Fehler bzw. Risiko und warum es relevant ist	Konkrete Korrektur oder Alternativtext
		typische generische Schlussformel.	45°C crossing between 20 and 25 minutes. The remaining uncertainty is large enough that I would use 20 minutes as a cautious reminder.
79	Quellen	Lienhard, Khan Academy und Abraham et al. stehen im Literaturverzeichnis, werden im Text aber nicht eindeutig zitiert. Das erschwert Nachvollziehbarkeit und kann wie Quellenpolsterung wirken.	Jede verwendete Quelle an der konkreten Aussage zitieren; unbenutzte Quellen entfernen.
80	Bild in Introduction	Das Hibiskusbild hat keine sichtbare Quellenangabe. „All experimental temperatures and graphs ...“ deckt ein fremdes Bild nicht ab.	Bild entfernen oder Urheber/URL/Lizenz direkt in einer Bildunterschrift und im Literaturverzeichnis angeben.
81	Works Cited	Zitierstil, DOI-/URL-Schreibweise und Abrufdaten sind nicht vollständig einheitlich.	Einen Stil durchgehend verwenden, z. B. MLA 9 oder APA 7; DOI möglichst als https://doi.org/....
82	OpenStax source	Der angegebene OpenStax-Link „4-key-equations“ verweist nicht eindeutig auf den beschriebenen Newton-Cooling-Abschnitt.	Exakten Abschnittstitel und direkte URL prüfen und zitieren.
83	Reproduzierbarkeit	„spreadsheet and graphing software“ nennt weder Software, Regressionsoption noch Formelbereich.	Software/Version und Vorgehen nennen, z. B. “I used ... to regress $\ln(T - T_a)$ on t and calculated RMSE with ...” ; zentrale Rechnung trotzdem im Text erklären.
84	Sprache	Mehrere Tippfehler schwächen den professionellen Eindruck: „seperated“, „linearisaton“, „Newtons law“, „degrees celsius“, teils fehlende Artikel und uneinheitliche Großschreibung.	Vollständige Rechtschreibprüfung; separated, linearisation, Newton's law, degrees Celsius.
85	Akademische Integrität	Wenn Formulierungen oder Rechnungen aus KI-Unterstützung in die endgültige IA übernommen werden, müssen sie nach der aktuellen IB-Regel transparent gekennzeichnet und referenziert werden. Ungekennzeichnete KI-Texte gelten nicht als eigene Arbeit.	Die mathematischen Ideen selbst nachvollziehen und in eigener Sprache neu verfassen. KI-generierte Bestandteile nur entsprechend den Vorgaben der Schule und des IB offenlegen; niemals den Bericht ungeprüft als eigenen Text einreichen.

5. Entscheidender mathematischer Neuaufbau

5.1 Ein Modell, mehrere Schätzmethoden

Die Arbeit sollte nicht A, B und C als drei physikalisch verschiedene Modelle darstellen. Die gemeinsame Modellannahme lautet

$$\frac{dT}{dt} = -k(T - T_a), k > 0,$$

mit der Lösung

$$T(t) = T_a + A e^{-kt}, A = T(0) - T_a.$$

Die Unterschiede liegen ausschließlich darin, wie A und k geschätzt werden:

1. Endpunktschätzung,
2. lineare Regression nach Logarithmierung,
3. willkürliche Zweipunktschätzung,
4. direkte nichtlineare Regression in der ursprünglichen Temperaturskala.

Die vierte Methode passt am besten zum gewählten Gütemaß RMSE, weil sie unmittelbar

$$SSE(A, k) = \sum_{i=1}^n \left[T_i - (T_a + A e^{-kt_i}) \right]^2$$

minimiert.

5.2 Kontrollrechnung mit den sichtbaren Mittelwerten

Bei festem $T_a = 22.4^\circ \text{C}$ liefert eine direkte nichtlineare Kleinste-Quadrate-Anpassung an die angegebenen Mittelwerte ungefähr

$$\hat{T}(t) = 22.4 + 58.87 e^{-0.04291t},$$

wobei t in Minuten gemessen wird. Daraus folgen

$$RMSE \approx 1.819^\circ \text{C}$$

und

$$t_{45} = -\frac{1}{0.04291} \ln \left(\frac{45 - 22.4}{58.87} \right) \approx 22.31 \text{ min.}$$

Diese Zahlen sind eine unabhängige Kontrollrechnung auf Basis der im PDF sichtbaren Werte. Sie müssen in der Originaldatei mit den ungerundeten Messdaten erneut berechnet und selbst nachvollzogen werden.

5.3 Fairer Vergleich der vorhandenen Gleichungen

Methode	RMSE, alle Werte	RMSE für $t \geq 10$ min	Vorhersage t_{45}
A: Endpunkte	$3.130^{\circ}C$	$3.243^{\circ}C$	19.28 min
B: korrigierte Log-Regression	$3.530^{\circ}C$	$1.604^{\circ}C$	21.30 min
C: Punkte 15/45	$7.214^{\circ}C$	$2.050^{\circ}C$	21.53 min
Direkte nichtlineare Regression	$1.819^{\circ}C$	$1.805^{\circ}C$	22.31 min

Damit ist C weder global noch auf $t \geq 10$ die beste vorhandene Schätzung. Noch wichtiger: Die beobachtete lineare Interpolation zwischen 20 und 25 Minuten ergibt pro Trial ungefähr

$$24.15, 23.75, 24.10 \text{ min},$$

also einen Mittelwert von etwa 24.00 min. Der Abstand zum nichtlinearen Modellwert von 22.31 min beträgt rund 1.69 min. Genau dieser Unterschied ist wertvolle Reflexion: Die Trial-Streuung ist klein, aber das einfache Ein-Exponential-Modell zeigt einen systematischen Modellfehler.

6. Konkreter Plan für eine Top-Bewertung

Priorität 1 - zwingend vor Abgabe

1. Faktor-10-Fehler in Gleichung C und Sekunden/Minuten-Fehler korrigieren.
2. Gleichung B mit genau einem Parametersatz neu erzeugen; Tabelle, Grafik und RMSE automatisch daraus ableiten.
3. Boyle's law vollständig entfernen und Newtons Modell korrekt erklären.
4. Alle Modelle auf derselben Datenmenge vergleichen.
5. Das falsche angepasste RMSE korrigieren und die Modellwahl neu schreiben.

Priorität 2 - stärkster Punktezuwachs

1. Direkte nichtlineare Regression ergänzen.
2. Residuenplot für das bevorzugte Modell zeigen.
3. Trial 3 als unabhängige Validierung verwenden oder einen vierten Trial erheben.

4. Schwellenzeit pro Trial und eine Streuung angeben.
5. Symbol- und Einheitentabelle einführen: t , T , T_a , A , k , r_i , RMSE.

Priorität 3 - von gut zu sehr gut

1. Arbeit um 3-5 Seiten kürzen, ohne Mathematik zu entfernen.
2. Persönliche Entscheidungen über echte mathematische Konsequenzen zeigen, nicht durch häufiges „I“.
3. Jeden Einschränkungsabsatz mit einer quantitativen Auswirkung verbinden.
4. Unbenutzte Quellen und das unzitierter Bild entfernen.
5. Alle Abbildungen, Tabellen, Seitenzahlen, Einheiten und Dezimalstellen vereinheitlichen.

7. Direkt nutzbare, natürlichere Alternativpassagen

Diese Textbausteine sind bewusst sachlich und weniger „rubrikhaft“. Sie sollten nicht ungeprüft übernommen, sondern verstanden und in die eigene Stimme übertragen werden.

Forschungsfrage und Bewertungsmaß

Research question. How well does Newton’s law of cooling describe the temperature of 300 mL of hibiscus tea in a room at approximately 22.4°C , and what time does the model predict for the tea to fall below my chosen threshold of 45°C ? I will assess the model using residuals, RMSE in degrees Celsius, and its error when predicting a trial that was not used to estimate the parameters.

Physikalische Modellannahme

I treated the room temperature as constant and assumed that the net cooling rate was proportional to the tea’s temperature difference from the room. With $T_a = 22.4^\circ\text{C}$ and $k > 0$, this gives $\frac{dT}{dt} = -k(T - T_a)$.

Because every recorded tea temperature was above T_a , solving the equation gives $T(t) = T_a + A e^{-kt}$, where A is the initial excess temperature and k is measured in min^{-1} .

Modellanpassung

Linearising the data is useful for checking whether an exponential pattern is plausible, but it changes the weighting of the errors. Since my final accuracy measure is RMSE in degrees Celsius, I also fitted

$T(t) = 22.4 + A e^{-kt}$ directly by minimising the squared residuals in the original temperature scale. This avoids choosing two convenient observations and uses every value in the selected training set.

Fairer Modellvergleich

I compared every parameter-estimation method on the same observations. On the restricted domain $t \geq 10$ min, the RMSE values for A, B and C were approximately 3.243°C , 1.604°C and 2.050°C respectively. Therefore C was not the most accurate of these three estimates on the domain relevant to the threshold. The direct nonlinear fit gave the lowest full-data RMSE, but its residuals still showed a systematic pattern, so the one-exponential model should not be treated as exact.

Schlussfolgerung

The direct nonlinear model predicts that the tea reaches 45°C after about 22.3 minutes. Linear interpolation of the two measurements surrounding the threshold gives about 24.0 minutes. The difference of roughly 1.7 minutes is larger than the variation between the three interpolated trial values, suggesting systematic model error rather than only random measurement variation. I would therefore report a practical crossing window of 22-24 minutes, with the raw observations guaranteeing only that the crossing occurred between 20 and 25 minutes at the present sampling interval.

Kritische Reflexion

The high R^2 value of the linearised graph initially made the exponential model appear stronger than it was. However, R^2 described the transformed variable $\ln(T - T_a)$, whereas my practical question concerns errors in degrees Celsius and minutes. Recalculating the fit in the original scale changed the preferred parameter estimates. This showed me that a visually straight transformed graph is evidence for a model form, but not proof that its predictions are the most accurate for the quantity I care about.

8. Quellen zur Bewertung

1. International Baccalaureate, [Mathematics: analysis and approaches subject brief - first assessments 2021](#).
2. International Baccalaureate, [Mathematics: analysis and approaches updates](#) - neuer Kurs ab August 2027, erste Prüfung Mai 2029.
3. International Baccalaureate, [Artificial intelligence in learning, teaching, and assessment](#).
4. International Baccalaureate, [Academic integrity](#).
5. International Baccalaureate, [Mathematics: analysis and approaches guide, first assessment 2021](#), öffentlich zugängliche Kopie mit den vollständigen Kriterien.

Abschließende Einschätzung

Die Arbeit hat großes Potenzial. Entscheidend sind **Konsistenz, faire Validierung und ehrliche quantitative Modellkritik**, nicht mehr Seiten oder kompliziertere Mathematik. Werden die fünf Hauptfehler behoben und direkte Regression, Residuen und Validierung sauber integriert, ist eine überzeugende Spitzen-IA möglich.