

12 単回帰（回帰の意味と散布図）

1. 回帰で「将来予測」を作る（判断ではなく予測）

今日やること

ハンバーガーショップのデータを使って、
「ポテトの注文数を増やすには、何が関係しているのか」を見つけ、
「このくらい売れそう」という予測を作ります。

- ・ 今日のキーワード：散布図・回帰直線・予測式 ($\hat{y} = ax + b$)
- ・ 回帰は、正しい／間違いを決める話ではなく、未来を予測する話
- ・ ゴールは計算ではなく、式の意味を言葉で説明できること

今日の到達目標

- ・ 散布図から「増える／減る／関係なし」を説明できる
- ・ 回帰直線と式を表示し、傾き a の意味を言える
- ・ 回帰式で「もし x がこうなら、 y はこのくらい」を予測できる

Slide01：第11回の立ち位置（運用）

【目的】検定（判断）から回帰（予測）へ、目的が変わることを先に固定する。

【口頭補足】回帰は「当てる魔法」ではなく「だいたいの予測モデル」。点は散らばるのが普通。

【実習仕様】前半は clean データ（外れ値なし）で基本を成功させる：散布図→直線→式→ R^2 →予測。

後半は outlier データ（外れ値あり）で「現実はズレる」を確認：メモ（品切れ／キャンペーン）と結びつける。

【運用】開始直後に一言確認：「今日は“判断”と“予測”的どっち？」→「予測」。

2. ストーリー：店長の次の目標は「ポテト売上UP」

前回まで店長は、**ポテトの重さのばらつきをデータで見て改善しました。**
（「本当に減ったと言ってよいか？」を判断できるようになった）

次の目標（現場の目標）

- ・ ポテトは**利益率が高い**。
- ・ だから店長は**ポテトの売上を伸ばすことを目標にしました。**

そこで店長は考えます：

- ・ 「お客様がハンバーガーと一緒にポテトを注文する数を増やしたい」
- ・ 「そのために、何が増えるとポテトも増えるのだろう？」

📌 今日はこの問い合わせに対して、データから予測の式を作る。

Slide02：ストーリーの狙い（例を変えずに目的だけ変える）

【目的】ポテト例を継続し、認知負荷を上げずに「回帰=予測」へ移る。

【口頭補足】前回まで：品質（ばらつき）を整える話。今回は：売上を伸ばす話。

【実習仕様】データは「日別」：ハンバーガー注文数（ x ）とポテト注文数（ y ）。まず clean で関係をつかむ。

【運用】学生に短く発言させる：「ポテトを増やすには何が大事？」 → 「客数／セット注文」など。

3. 前回の検定は「判断」／今回の回帰は「予測」

前回までやってきたのは、次のような判断です。

「平均との差がある」だけではなく、ばらつきの中で珍しいかで決めました。

検定（前回）＝判断

珍しさ（p値）を使って、
「偶然と言ってよいか／言いにくいか」を決める。

一方、今日は次の目的です。

回帰（今日）＝予測

2つの数値の関係を使って、
「もし x がこうなら、 y はこのくらい」を予測する。

- ・ 検定は「棄却／棄却できない」で結論を出す
- ・ 回帰は「予測の式」を作り、使う（結論の形が違う）

Slide03：混同防止（今日のモード切替）

【目的】検定（判断）と回帰（予測）を混ぜない。結論の形が違うことを固定する。

【口頭補足】今日もデータはばらつく。回帰直線は「全部当たる線」ではなく「だいたいの中心」。

【実習仕様】clean では回帰の基本（式と R^2 、予測）を成功させる。outlier では 1 点の影響とメモ（品切れ／キャンペーン）で理由づけする。

【運用】ミニ確認：検定の結論は？ → 「棄却／棄却できない」。回帰の成果物は？ → 「予測の式」。

4. 問いの定式化：ハンバーガーが増えるとポテトは増える？ ($x \rightarrow y$)

店長の目標は「ポテト売上 UP」でした。

そこで、**ポテトの注文数**に関係しそうなものを考えます。

今日の問い合わせ (回帰の問い合わせ)

ハンバーガー注文数が増える日ほど、ポテト注文数も増えるのか？

回帰では、2つの数値の役割を先に決めます。

- ・ x (説明変数) : ハンバーガー注文数 (原因 かもしれない 側)
- ・ y (目的変数) : ポテト注文数 (増やしたい結果の側)

📌 今日やることは、 x を見たら y を予測できるかを確かめること。

Slide04：問い合わせを数学の形に落とす（混乱防止）

【目的】「何を予測したいか」を1行で固定し、 x と y の役割を先に決める。

【口頭補足】回帰は「 x が増えると y が増えるかもしれない」をデータで見る。ここでは因果は断定しない。

【実習仕様】clean/outlierとともに、 x =ハンバーガー注文数、 y =ポテト注文数で統一する（途中で変えない）。

【運用】学生に確認：「今日の y は何？」→「ポテト注文数」。

5. 散布図とは：2つの数値の関係を「点」で見る図（定義）

回帰を始める前に、まず散布図で関係を目で見ます。

散布図（定義）

1つのデータ（1日）を1つの点として、
横軸に x 、縦軸に y をとって並べた図。

今回の例では、点の意味はこうです。

- ・ある1日：(ハンバーガー注文数 x , ポテト注文数 y)
- ・30日分なら、点が30個できる

📌 散布図を見ると、増える／減る／関係なしが直感的に分かる。

Slide05：散布図を先に固定する理由

【目的】回帰直線の前に「点の雲」を理解させる（線の意味が分かる土台）。

【口頭補足】散布図は「同時に観測した2つの数値」の関係を見る図。1点=1日の記録。

【実習仕様】まず散布図だけ作らせる（近似直線は次段階）。軸ラベルと単位も入れる。

【運用】ここで小確認：「1点は何？」→「1日のデータ」。

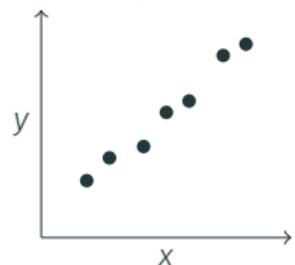
6. 散布図の読み方①：右上がり／右下がり／関係なし（3パターン）

散布図は、点の並び方で「関係の方向」を読みます。

（細かな計算の前に、まず目で判断できるようにする）

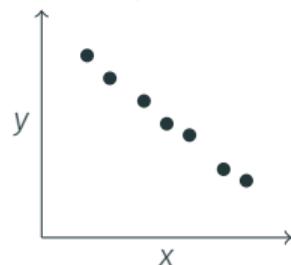
右上がり

（ x が増えると y も増えやすい）



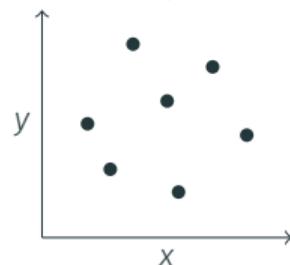
右下がり

（ x が増えると y は減りやすい）



関係なし

（ x が変わっても y が読めない）



📌 今日のデータ（ハンバーガー → ポテト）は、右上がりになりそうかをまず確認する。

Slide06：模式図を入れる理由（視覚で固定）

【目的】留学生でも「方向」を一目で判断できるように、まず3パターンを固定する。

【口頭補足】この段階では「強さ」はまだ言わない。「増える／減る／読めない」の3分類だけ。

【実習仕様】散布図を作ったら、まずこの3つのどれに近いかを言語化させる（1行でOK）。

【運用】学生に挙手で答えさせる：「今日の関係はどれっぽい？」

7. 散布図の読み方②：ばらつき・外れ値（現実データの特徴）

現実のデータは、きれいに一直線には並びません。
同じ x でも y が少し違うのが普通です。

重要：データには「ばらつき」がある

ばらつきがあるから、予測は 100% 当たるわけではない。
でも、中心の傾向をつかめば、だいたいの予測はできる。

さらに、外れ値が混ざることもあります。

- ・ 例：品切れでポテトが売れなかった日（普段より極端に少ない）
- ・ 例：キャンペーンでポテトが急に売れた日（普段より極端に多い）

➡ 後半の実習では、外れ値があると回帰直線がどう変わるかを確かめる。

Slide07：ばらつきと外れ値を先に言う（後半実習の伏線）

【目的】「回帰直線＝全部当たる線」という誤解をここで止める。

【口頭補足】ばらつきは悪いものではない（現実そのもの）。外れ値は「理由」があることが多い。

【実習仕様】前半 clean で基本を成功（外れ値なし）。後半 outlier で外れ値の影響を見る。メモ列で理由づけする（品切れ／キャンペーン）。

【運用】外れ値を見つけたら「なぜ？」をメモで確認させる（説明できればOK）。

8. 回帰の目的：点の雲を「一本の線」で要約し、予測に使う

散布図を作ると、点は雲（ばらついた集まり）になります。
このままだと、「次はどれくらい？」が言いにくい。

回帰の目的（理由）

点の雲を**一本の線**で要約して、
予測に使える形（式）にする。

回帰でやりたいことは、次の2つです。

- ・ 関係を要約する： x が増えると y はどう変わりやすいか
- ・ 予測する：もし x がこの値なら、 y はこのくらい (\hat{y})

📌 ここで「線」は、全部当てる線ではなく、中心の傾向を表す線。

Slide08：回帰が必要な理由（散布図→線への納得）

【目的】「なぜ線を引くのか」を言葉で納得させる（式の前に理由）。

【口頭補足】点がバラバラでも、中心の流れが見えるなら予測はできる。回帰はその中心を“一本化”する。

【実習仕様】clean で「線を引くと予測できる」成功体験を作る。outlier で「線がぶれる」も体験させる。

【運用】学生に一言で言わせる：「回帰の目的は？」 → 「点を線で要約して予測する」。

9. 回帰直線（イメージ）：各 x に対して “中心” を通る予測の線

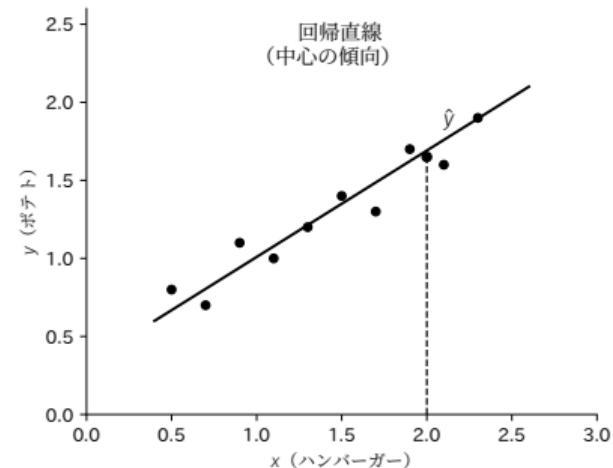
回帰では、点の雲の中心を通るような線を考えます。
この線を回帰直線と呼びます。

回帰直線（イメージ）

同じ x のときの y の “中心” を結んだような
線。
この線上の値が、 y の予測値 (\hat{y})。

- 点は線の周りに散らばる（ばらつき）
- それでも線があると、だいたいの予測ができる

📌 点の中心を表す線だと思えばよい（まずはこれでOK）。



点が散らばっていても、線があれば
「この x なら \hat{y} はこのくらい」が言える

Slide09：回帰直線の直感（数式前に図で固定）

【目的】「回帰直線＝中心の傾向」「 \hat{y} = 線上の予測値」を図で固定する。

【口頭補足】同じ x でも y は毎回違う。その中心を線で表す。点が線からずれる分が“誤差/残差”。

【実習仕様】Excel では散布図→近似直線を表示し、線上の式が出ることを確認させる。

【運用】学生に指さし確認：「 \hat{y} はどこ？」 → 「線の上」。

10. 回帰式： $\hat{y} = ax + b$ (\hat{y} = 予測値) を文章で理解する

回帰直線は、式で表せると便利です。

(Excel が計算して表示してくれる式を、言葉で説明できるのが今日の目標)

$$\hat{y} = ax + b$$

この式が言っていること（文章）

x (ハンバーガー注文数) が分かれば、
ポテト注文数の予測値 \hat{y} を計算できる。

- ・ \hat{y} (ハット付き) は予測値 (実測の y とは違う)
- ・ 点はばらつくので、実測 y は \hat{y} と一致しないことが多い

📌 今日の実習は、Excel で式を出し、その意味を言葉で説明すること。

Slide10：式を“暗記”させないための言い換え

【目的】 $\hat{y} = ax + b$ を数式としてではなく「予測の計算ルール」として理解させる。

【口頭補足】 \hat{y} は“当たった値”ではなく“予測した中心”。ハットが付くのは「予測だから」。

【実習仕様】近似直線の式を表示し、 x に具体的な数（例：150）を入れて \hat{y} を計算させる。

【運用】短い確認：「 y と \hat{y} の違いは？」 \rightarrow 「 y は実測、 \hat{y} は予測」。

11. パラメータの意味：傾き a = 増え方／切片 b = $x = 0$ の予測

回帰式 $\hat{y} = ax + b$ の意味は、 a と b を見れば説明できます。

傾き a (いちばん大事)

x が 1 増えると、 \hat{y} が平均でどれくらい増えるか。

例： $a = 0.6$ なら、バーガーが 10 増えるとポテトは約 6 増える。

切片 b (注意して使う)

$x = 0$ のときの予測値。

ただし、現実に $x = 0$ (バーガー 0) が起きない範囲なら、
 b の現実的な意味は薄いことがある。

📌 実習では、まず a (増え方) を言葉で説明できることを優先する。

Slide11 : a を主役、b は最小限（混乱防止）

【目的】初心者がつまずきやすい「切片の意味」を最小限にし、傾きの解釈を主役にする。

【口頭補足】ビジネス上の意味はたいてい傾き（増え方）。切片は“計算の都合”として出ることもある。

【実習仕様】clean データの式で、傾き a を「10 増えたら何増える？」に言い換えさせる。

【運用】学生に口頭で言わせる：「a は何？」 → 「増え方（ x が 1 増えたときの y の増え方）」。

12. R^2 とは：線で説明できる割合の目安（正解率ではない）

回帰直線を引いたら、次に気になるのはこうです。

「この線は、どれくらいデータに合っているの？」

R^2 （決定係数）

回帰直線で、 y のばらつきのうち
どれくらいが x で説明できたかを表す目安。

- R^2 は 0~1 の間の数（例：0.72）
- 1 に近いほど、点が線に近く、説明できる割合が高い
- 0 に近いほど、点が線から散らばり、説明できる割合が低い

ここが重要（誤解防止）

R^2 は 正解率ではない。

「未来が必ず当たる」ことは保証しない。

Slide12 : R^2 の言い方を固定する (留学生対応)

【目的】 R^2 を「当てた回数」や「正解率」と誤解させない。意味を 1 文で固定する。

【口頭補足】 R^2 は「線で説明できる割合の目安」。未来が当たる保証ではない。

【実習仕様】 Excel の近似直線オプションで「 R^2 値を表示」を ON にし、 clean と outlier で値がどう変わるか比較させる。

【運用】 学生に一言で言わせる：「 R^2 って何？」 → 「線で説明できる割合の目安」。

13. 注意：相関≠因果／回帰は「当てる保証」ではない

回帰ができると、つい次のように言いたくなります。

「バーガーが増えたから、ポテトが増えた！」

でも注意

散布図や回帰で分かるのは、基本的に**関係（相関）**であって、
原因（因果）を確定することではない。

例：バーガーとポテトが同時に増える本当の理由は、別にあるかもしれません。

- ・ その日は**来客数**が多かった（両方増える）
- ・ キャンペーンをやっていた（セット注文が増える）
- ・ 品切れがあった（片方だけ減る）

今日の結論の形

回帰で言えるのは、

「この範囲では、増える傾向がある／予測に使えるかもしれない」まで。

Slide13：相関と因果の線引き（やりすぎない）

【目的】回帰=因果の証明、と誤解させない。特に留学生は断定表現に引っ張られやすい。

【口頭補足】回帰は「予測の道具」。原因は別調査（実験・追加データ）が必要になることが多い。

【実習仕様】outlier の回で、メモ列（品切れ／キャンペーン）を使い「別要因がある」ことを体験させる。

【運用】ここでは深掘りしない。「相関≠因果」を合言葉として1回言って終える。

14. 実習で使うデータ：clean / outlier の 2 つ（目的が違う）

今日は同じテーマ（ポテト売上 UP）でも、目的の違う 2 種類のデータで学びます。
順番が大事です（まず成功→次に現実）。

データ A : clean (外れ値なし)

回帰の基本を「成功」させるためのデータ。
散布図が読みやすく、回帰直線がきれいに出る。

- ・ 目的：散布図→直線→式→ R^2 →予測、をスムーズに体験
- ・ 「傾き a の意味」を言葉で説明できるようにする

データ B : outlier (外れ値あり)

現実のデータはズレることを体験するためのデータ。
品切れ／キャンペーンなどで点が極端に外れる日が混ざる。

- ・ 目的：外れ値が回帰直線・ R^2 ・予測に与える影響を知る
- ・ 「なぜズレたか」をメモ列から説明する

Slide14：2 データ構成の狙い（疲れない授業設計）

【目的】回帰の基本を“成功体験”で固めてから、現実（外れ値）に進む流れを宣言する。

【口頭補足】同じ回帰でも、学ぶ目的が違う。clean は「型を覚える」、outlier は「現実の注意点」。

【実習仕様】Excel ファイル内でシートを分ける（例：clean / outlier）。列構成は同じにして比較しやすくする。

【運用】学生に宣言：「まず A で回帰を完成させる。次に B で“なぜズレる？”を考える」。

15. 実習のゴール：散布図→直線→式→R²→予測→（外れ値でズレを考える）

実習では、回帰を「作って使う」までを一気にやります。
最後に、外れ値で現実のズレも確認します。

実習の流れ（この順番で固定）

1. 散布図を作る (x と y の関係を点で見る)
2. 回帰直線を引く (点の雲を一本の線で要約)
3. 回帰式 $\hat{y} = ax + b$ を表示する (\hat{y} = 予測値)
4. R² を表示する (線で説明できる割合の目安)
5. 具体的な x を入れて 予測する (\hat{y} を計算)
6. (後半) outlier で ズレの原因を考える (メモ列と結びつける)



「傾き a は何を意味する？」／「ズレた日はなぜ？」

Slide15：成果物の形（評価ポイントの明確化）

【目的】実習のゴールを「操作」ではなく「説明できる成果物」に置く。

【口頭補足】計算は Excel がする。大事なのは“意味を言える”こと (a 、 \hat{y} 、 R^2 、外れ値)。

【実習仕様】確認物：①散布図（軸ラベルあり）②式と R^2 表示 ③予測例（ x の値と \hat{y} ）④外れ値の説明（メモ参照）。

【運用】時間が厳しい場合：clean で⑤まで必達、outlier は⑥を口頭共有だけでも可。

16. 実習①-1 (clean) : 散布図を作る (x =バーガー数、 y =ポテト数)

目的：回帰はまず散布図から始める。点の並びで「関係の方向」を読む。

やること（操作）

1. clean シートを開く（外れ値なし）
2. x （横軸）に バーガー注文数、 y （縦軸）に ポテト注文数 を選択
3. 散布図（点）を作る

必ずやる（見た目のルール）

- ・ グラフタイトル：「バーガー数とポテト数 (clean)」
- ・ 横軸ラベル：バーガー注文数（件）
- ・ 縦軸ラベル：ポテト注文数（件）

📌 まずは右上がり／右下がり／関係なしのどれに近いかを言えるようにする。

Slide16：実習仕様（授業内の確認ポイント）

【目的】回帰の出発点は散布図。軸の意味を固定してから線に進む。

【口頭補足】点が右上がりなら「増える傾向」。この段階では“強さ”は言わない。

【実習仕様】提出なし。教師は教室を回り、(1) 散布図になっているか (2) 軸ラベルが正しいかを確認。

【運用】早い学生に質問：「この散布図は3パターンのどれ？」→一言で答えさせる。

17. 実習①-2 (clean) : 回帰直線・式・R² を表示 (線形近似)

目的：点の雲を一本の線で要約し、式と R² を表示する。

やること（操作）

1. 散布図上の点をクリック
2. 近似曲線（トレンドライン）を追加（線形を選ぶ）
3. 「グラフに数式を表示する」を ON
4. 「グラフに R² 値を表示する」を ON

表示場所の指示（読みやすくする）

- ・ 式と R² は、点に重ならない場所へドラッグして移動
- ・ 点が見えにくい場合は、グラフを少し拡大して OK



ここで得るものは2つ：回帰式（予測ルール）と R²（当てはまりの目安）。

Slide17：Excel 操作で迷わせない（最短手順）

【目的】操作で詰まると回帰の意味が飛ぶため、手順を短く固定する。

【口頭補足】線は“中心の傾向”。式は“予測の計算ルール”。 R^2 は“説明できる割合の目安”。

【実習仕様】提出なし。教師は(1)線形近似になっているか(2)式と R^2 が表示されているかを確認。

【運用】うまく出ない学生への定番支援：「点を選択→近似曲線→線形→式表示→ R^2 表示」。

18. 実習①-3 (clean) : 読み取りと予測 (傾きの解釈 + 指定 x で \hat{y} 計算)

目的：出てきた式を「読める」「使える」にする。
(計算よりも意味を言葉で説明できることがゴール)

Step1: 傾き a を言葉で説明する

回帰式 $\hat{y} = ax + b$ の a は、
バーガーが 1 増えたとき、ポテトが平均でどれくらい増えるか。

- 例： $a = 0.6$ なら「バーガーが 10 増えると、ポテトは約 6 増える」

Step2: 指定 x で予測値 \hat{y} を計算する

教師が指定する x (例: $x = 150$) を式に代入し、 \hat{y} を計算する。
線の上の値 = 予測値 (ハット付き)。

📌 口頭チェック: y は実測、 \hat{y} は予測 (同じではない)。

Slide18：授業内の確認（提出なし運用）

【目的】回帰の本質（傾きの解釈+予測）を“言葉”で言えるようにする。

【口頭補足】 b （切片）は今回は深入りしない。主役は a （増え方）。 \hat{y} は中心の予測で、当たりを保証しない。

【実習仕様】提出なし。教師指定の x を黒板/スライドで提示し、学生は各自の式で \hat{y} を計算。

【運用】数名を指名して発表：① a の意味を日本語で言う ② $x = 150$ のときの \hat{y} を言う（値は多少ズレてもOK、式の違いが原因と説明）。

19. 実習② (outlier) : 外れ値入りで再実行 (線・式・R² の変化を見る)

目的：現実のデータは「きれい」ではない。

外れ値が入ると、回帰直線・式・R²・予測がどう変わるかを体験する。

● やること (操作は実習①と同じ)

1. outlier シートを開く
2. 散布図を作る (軸ラベルも同じ)
3. 線形の回帰直線を追加し、式と R² を表示する

● 比較するポイント (ここが学び)

- ・ 線の向き・位置は変わったか？
- ・ 傾き a はどれくらい変わったか？
- ・ R² は上がった？ 下がった？
- ・ 同じ x (例: $x = 150$) で \hat{y} はどれくらい変わる？

● 理由づけ: メモ列で説明する

- ・ 外れている点を見つけたら、メモを見て理由を言葉にする。

例: 品切れ (本当は売れるのに売れない) / キャンペーン (急に増える)

Slide19：外れ値を“発見→説明”させる運用

【目的】外れ値の影響（線・式・ R^2 ・予測のズレ）を、計算ではなく比較で理解させる。

【口頭補足】外れ値は「ミス」ではなく「現実の事情」を含むことがある。メモはその手がかり。

【実習仕様】提出なし。教師指定で「外れている点を1つ見つけて、メモで理由を言う」を必達にする。

【運用】共有のしかた：2～3名に発表させる（①どの点が外れ値？②メモは？③それで線や R^2 はどう変わった？）。

20. まとめ

今日の到達点（ここだけ）

- 散布図で「関係」を見て、回帰直線で一本の線に要約した
- 回帰式 $\hat{y} = ax + b$ を使って、予測を計算できるようになった
- R^2 は「当てはまりの目安」で、正解率ではないと分かった

現実のポイント：ズレ（残差）が必ずある

実測 y と予測 \hat{y} は一致しないことが多い。

このズレを 残差（誤差）と呼ぶ。

- clean：ズレは小さめ → 回帰の基本が見える
- outlier：ズレが大きくなる → 予測がぶれる／ R^2 も変わる

次回への：では、線はどうやって決めている？

Excel は勝手に線を引いているように見えるが、

実は「ズレ（残差）を小さくする」ルールで線を決めている。

次回はその決め方＝最小二乗法を学ぶ。

Slide20：次回（最小二乗法）への接続を自然にする

【目的】回帰の“結果”から“決め方”へ関心を移す。最小二乗法は「ズレを小さくするルール」として導入。

【口頭補足】残差=点から線までの縦の距離（まずはこの直感でOK）。なぜ二乗？は次回に回す。

【実習仕様】最後に口頭確認：①回帰は何を作る？→「予測モデル」②ズレの名前は？→「残差」。

【運用】次回の導入に使える一言問い合わせ：「線はどうやって決めたの？」→「ズレを最小にする（最小二乗法）」。