# Journey to the Stratosphere 



## Introduction


https://bit.Iy/LinearRegression30XPlus

http://bit.ly/HighAltitudeBalloon


Stratosphere Video

## Teacher Notes:

This video is amazing. Watch as the sky changes from a pale blue to almost black. Listen to the changes as the atmosphere becomes less dense. Watch the balloon as it increases in size, so much to observe. Many schools around the world have actually done this experient. But, if you don't have the time or a spare couple of thousand dollars, watch the video.

This activity focuses on the relationship between time and altitude. "Does the balloon continue hurting upwards at the same rate?" Posing this question at the start of the


Calculator Instructions activity and engaging students in a discussion generates a level of curiosity and intellectual 'buy-in'. To add further interest, the video can continue playing as students work through the questions with a view to answering Question 14 (predictions) "What will be the height of the balloon after 30 minutes" before the balloon actually reaches that point. Playing the video during the activity perpetuates the context and creates a level of aniticipation. If students have bought into the discussion, building the suspence will elicit dopamine which will increase the likelihood of students remembering the lesson. In addition to altitude, pressure and temperature can also be investigated.

Release a helium filled balloon and watch it race skyward. How high will it go? Where will it end up? Does it continue hurtling upward at the same rate that it left your hand? Maxime Dehaye Web Developer and avid photographer released a weather balloon with a payload consisting of a bundle of telemetric devices and some cameras. The result has been uploaded to YouTube for all to see and share. In this investigation you will use the first two minutes of data from Dehaye's video to predict the altitude of the balloon some 30 minutes later, well beyond extrapolation norms.

A series of videos to assist with data entry, manipulation and regression analysis have been created and available on YouTube. Scan the QR codes opposite to view the videos.

## Collecting the Data

Watch Maxime's video on YouTube and record the time (10 second intervals) and corresponding altitude for the first two minutes of the balloon flight.

Store the time data in List $1\left(\mathrm{~L}_{1}\right)$ and the altitude data in List $2\left(\mathrm{~L}_{2}\right)$.

| 4 | [\| DEG | 困 |
| :---: | :---: | :---: |
| 0 | 127 |  |
| 10 | 169 |  |
| 20 | 223 |  |
| 30 | 275 |  |
| L1 |  |  |

Data entry can be accessed by pressing the [DATA] key.

## Question: 1.

At what altitude does the balloon flight commence?
Answer: 127 m

## Question: 2.

Study the altitude values, in particular, the difference between consecutive altitudes.
i) What is the 'average' change in the altitude in each 10 second interval?

Answer: Consecutive differences: 42, 54, 52, 46, 50, 55, 43, 50, 48, 51, 46, 50. Mean: 48.9
ii) What is the 'average' change in the altitude per second?

Answer: 4.89 m/s

## Question: 3.

Based on your answers from Questions 1 and 2; suggest a possible straight line equation to model the data.
Answer: Altitude $=4.89 \mathrm{~T}+127 .[y=4.89 x+127]$

## Equation Quality

In this section we explore how well your equation models or predicts the altitude of the balloon.

## Question: 4.

Graph the altitude of the balloon against time and include your straight line graph from Question 3.


[^0]Author: P. Fox

## Question: 5.

Calculate the average (mean) time and altitude for each data set and plot this as a point on your graph. Answer: Mean time: mean $\left(\mathrm{L}_{1}\right)=60$. Mean Altitude: mean $\left(\mathrm{L}_{2}\right)=420.23$ [Refer Graph]

## Question: 6.

Does your straight line pass through the point plotted in Question 5?
Answer: Students may use visual inspection to see if the line passes through the graph. Due to the accuracy of the prediction and the 'estimations' pertaining to graphing by hand, students may conclude either way as to whether or not the graph passes through $(\bar{x}, \bar{y})$. Based on the proximity students should seek a more accurate result by performing the calculations via substitution: $420 \approx 4.89 \times(60)+127$.

The equation is close since $4.8916 \times(60)+127=420.5$, but the ultimate conclusion is that the estimated line does not pass through: $(\bar{x}, \bar{y})$.

The straight line graph passes quite close to each point. The distance from each point can be measured or quantified by comparing predicted values with the data gleaned from the video. On the calculator this can be done very efficiently.

The first step is to define your rule.
Press the [table] key then select option 1: Add / Edit Function.


Type in the equation that you derived in Question 3.
Use the variable key $[x, y, z, t \ldots]$ to access ' $x$ '.
Once the rule or function has been defined press [ 2nd ] [ mode ]
Now it is time to enter this formula into a list.
Press the [ data ] key and navigate across to list 3: $L_{3}$.

Press the [ data ] key again and navigate across to the FORMULA menu, select option 1: Add / Edit Formula.


To use the rule or function to calculate the predicted values, press [ table ] and select option 2: f(

Press the [ data ] key to access the list names and select List 1: $L_{1}$.


Close the parenthesis to complete the formula: $f\left(L_{1}\right)$ and press [ enter] to generate the set of predictions in List 3.

[Predicted values will populate List 3 after pressing enter]

## Question: 7.

Determine the differences between the predicted values (using your rule) and the actual data.
Answer: Numerical differences: $0 ., 6.916,1.832,-1.252,1.664,0.58,-5.504,0.412,-0.672,0.244,-1.84$, 1.076, and -0.008

The formula in List 3 can be modified to calculate the numerical differences between the predicted value and the actual data. Formula: $f\left(L_{1}\right)-L_{2}$

## Question: 8.

Determine the average (mean) of the numerical differences and suggest reasons why this calculation is not a true representation of the 'average' distance.
Answer: Mean of numerical distances $(Q 7)=0.2652$. The combination of positive and negative numbers can generate an average close to zero which is not representative of the 'distances'.

## Question: 9.

Determine sum and mean of the squared differences.
Answer: The sum of the squared differences $=91.378$ and the mean of these squared values $=7.029$

The formula in List 3 can be edited to calculate the squared differences between the predicted value and the actual data. Formula: $\left(f\left(\mathrm{~L}_{1}\right)-\mathrm{L}_{2}\right)^{2}$. To 'insert' characters press [ 2nd ] followed by [ delete ]. The data > statistics menu can be used to calculate the sum and the mean.

## Least Squares Regression

One numerical approach for obtaining a 'line of best' fit is called Least Squares Regression. As the name suggests the equation is determined by generating the minimum sum of the square areas drawn such that one vertex of the square is on the line and another through the data point.

The Least Squares Regression line can be computed directly by the calculator.

Press [ 2nd ] [ data ] and navigate to option 4: Linear Regression.

Match the settings shown opposite, remembering that:

$$
\begin{aligned}
& \mathrm{L}_{1}=\text { Time } \\
& \mathrm{L}_{2}=\text { Altitude }
\end{aligned}
$$

The regression equation will be stored in $\mathrm{g}(\mathrm{x})$. Select Calculate!


## Question: 10.

Write down the Least Squares Regression equation generated by the calculator and include the $r^{2}$ value. The calculator stores the equation in the form: $y=a x+b$.

Answer: $y=4.9159 x+125.27 r^{2}=0.9998$

## Question: 11.

Determine the sum of the squared differences using this new equation. You can use the steps provided previously and changing $f(x)$ to $g(x)$ in the formula.
Comment on the result and compare with the value obtained in Question 9.
Answer: Sum of squared differences $=79.687$. This value is less than the previous calculation. (Better fit)

## Question: 12.

Does the least squares regression line pass through the point: $(\bar{x}, \bar{y})$ ?
Answer: Yes. The mean value of the altitudes is 420.2307 and the Least Squares Regression line passes through the point $(60,420.2307)$. Note that students can use $g(60)$ to compute this value.

## Question: 13.

Which line is a better fit? The original line determined in Questions $1-3$ or the Least Squares Regression line?
Answer: Whist the two lines in this investigation are very close the Least Squares regression line is the better fit as the average distance from the line to each data point is less. (Mean of squared distances)

## Question: 14.

Use your original equation and the Least Squares Regression equation to calculate the altitude of the weather balloon at the following times:
i) $\quad 5$ minutes ( 300 seconds)

Answer: Original Eqn: $f(300)=1594 \mathrm{~m}$ Least Squares: $g(300)=1600 \mathrm{~m} \quad$ Actual $=1677 \mathrm{~m}$ (4.8\%)
ii) 10 minutes

Answer: Original Eqn: $f(600)=3061 m \quad$ Least Squares: $g(600)=3704 m \quad$ Actual $=3153 m(2.4 \%)$
iii) 30 minutes

Answer: Original Eqn: $f(1800)=8931 \mathrm{~m}$ Least Squares: $g(1800)=8973 \mathrm{~m}$ Actual $=9161 \mathrm{~m}(2.1 \%)$
iv) 2 hours

Answer: Original Eqn: $f(7200)=35346 \mathrm{~m}$ Least Squares: $g(7200)=35520 \mathrm{~m}$ Actual $=127 \mathrm{~m}$
At an altitude of approximately $23,300 \mathrm{~m}$ ( 1 hr 27 m into flight) the weather balloon pops! At 2 hrs the balloon has just arrived back at ground level.

[^1]In each case above, comment on the accuracy of the prediction.

## Teacher Note:

Due to the remarkable linearity of the data, the predicted value at 30 minutes is very accurate. It should be noted that it is not common practice to extrapolate so far beyond a regression line. In a sense Part (iv) is designed to reinforce that notion.

## Least Squares Regression - Extension

## Teacher Note:

Whilst this section may not be specifically on the applicable mathematics course, it is a simple and brief inclusion and illustrating how so much of statistics can be connected. The idea of the activity goes beyond 'pressing buttons' to generate a Least Squares Regression equation, it is designed to generate understanding and engagement in statistics.

## Question: 15.

Use the statistical features of the calculator to determine the following:
i) Sample standard deviation of the time data. $\left(\mathrm{S}_{\mathrm{T}}\right)$ Answer: 38.944
ii) Sample standard deviation of the altitude data. ( $\mathrm{S}_{\mathrm{A}}$ ) Answer: 191.465
iii) Pearson's correlation coefficient for the time - altitude data. (r) Answer: 0.999909

## Question: 16.

Use your answers from Question 15 to calculate: $r \frac{S_{A}}{s_{T}}$ and compare the result with the gradient of the Least Squares Regression equation.
Answer: $r \frac{s_{A}}{s_{T}}=0.9999 \times \frac{191.465}{38.944} \approx 4.91597$. Gradient of Least Squares Regression Equation: 4.9159

## Question: 17.

Assuming the Least Squares Regression equation passes through the point: $(\bar{x}, \bar{y})$, use your answer to Question 16 to determine the Least Squares Regression equation.
Answer: $420.2307=4.9159 \times 60+C . \therefore C=125.28$ Least Squares Regression Eqn: $y=4.9159 \mathrm{x}+125.28$


[^0]:    (C) Texas Instruments 2020. You may copy, communicate and modify this material for non-commercial educational purposes provided all acknowledgements associated with this material are maintained.

[^1]:    (c) Texas Instruments 2020. You may copy, communicate and modify this material for non-commercial educational purposes

