

Olympic Records

Issues

Summer Olympics 2020 will soon be here. Olympians and coaches often look at data from past world record events to inform their training. There is no single, simple formula that predicts what the best performance of male and female athletes should be. This is partly because the physiological attributes required to be a good sprinter, for example, are quite different from those required to be a good long distance runner. There are also many factors that go into determining who is the best runner at a particular distance, some genetic, some to do with training, some to do with where you have lived and trained during your life. Faced with this complicated collection of factors, it is useful to look at the data to see if there are any strong patterns in it.

- Is there a correlation between two different factors?
- Are there any simple trends about past performance that we can use to predict what might happen in the future?

Objective

Students will research world record information to mathematically model, process, and analyze the data. Students will demonstrate their understanding of scatterplots, correlation, and lines of best fit and use these tools to make predictions.

Procedure

Choose whether you wish to explore running, swimming, or speed skating. Look at the world record times for the events as broken out below.

Flat running events : 100m, 200m, 400m, 800m, 1500m, 1000m, 2000m, 3000m, 10000m, marathon *Freestyle swimming events* : 50m, 100m, 200m, 400m, 800m, 1500m *Speed skating events* : 500m, 1000m, 1500m, 3000m, 5000m, 10000m

Do this first for the men's world records and then for the women's world records.

1. Work out the average speed that is run/swam for each distance.
2. Create a Google Sheets file to house this table. Make the first column to be your labels (1000 meter, etc.), then distance (x), and third column average speed (y).
3. Make Sheet1 be for the men's data and Sheet2 for women's.
4. Create a Starter File in Pyret that retrieves data from this Google sheet. You may find it helpful to watch <https://www.youtube.com/watch?v=ltEgqIOROTU> to get started!
5. Use Pyret to create scatterplots that plot average speed against distance for both the men's records and the women's records.

For both the men's & the women's graphs

1. Describe the shape of the graph.
2. Add a line of best fit to the plot and record its equation.
3. Interpret the slope in context of distance and average speed.
4. Interpret in context the y-intercept of the line of best fit.
5. Interpret the correlation coefficient (r) in context.
6. Interpret the coefficient of determination (r^2) in context.
7. Compare the men's events to the women's events.
8. Using one of your two scatterplots, create a problem that asks for a prediction using the line of best fit and calculate the prediction. How confident are you to use this prediction to assist the athletes?

Produce

1. Create a written report that includes:

2. Calculations for the average speed
3. Scatterplots and shape descriptions
4. Scatterplots with line of best fit graphed, along with equation
5. Answers to numbers 3-8 above
6. Attach a page of all Pyret code used, along with how you set up your starter file

Create one slide that summarizes your analysis and *another slide* for your prediction question, including the scatterplot you are referencing. You will ask the class to make the prediction. Be prepared to let them know if they predicted correctly from your graph.

As a class we will analyze how the different sports compare.

Resources

Wikipedia World Record progressions for [running](#)

Wikipedia World Record progressions for [swimming](#)

Wikipedia World Record progressions for [speed skating](#)

Correlation Project Rubric

Name:	Got it!	Mostly	Getting There	Somewhat	Nope
Scatterplots: Calculations for average speed were shown and correct. Both scatterplots were created with and without line of best fit. Shape of graph was described.	35	27	21	10	0
Analysis: Line of Best Fit was correctly interpreted, including slope, y-intercept, r, and r^2 descriptions in context. Men's and women's records were compared using data science language. Prediction question was created and answered correctly.	40	31	24	12	0
Pyret Code: All coding used in Pyret was attached and correct, including the start file for the table.	15	13	11	4	0
Product: Project is neat, organized. Clear that effort was put in. Both a paper and summary slides were submitted on time.	10	9	6	3	0

(Project designed by: Joy Straub)