

Metodi matematici e simulazione numeriche

Regressione lineare per l'Intelligenza Artificiale

Cinzia Elia @ Orientamento Consapevole, 17 Aprile 2023
Dipartimento di Matematica



Cos'è l'Intelligenza Artificiale?

Cos'è l'Intelligenza Artificiale?

Capacità dei computer di imitare funzioni cognitive della mente umana quali l'apprendimento e la soluzione di problemi.

Cos'è l'Intelligenza Artificiale?

Capacità dei computer di imitare funzioni cognitive della mente umana quali l'apprendimento e la soluzione di problemi.

Applicazioni: riconoscimento del linguaggio scritto e parlato (es. chatbot, ChatGPT), predizioni (es. mercato finanziario), guida autonoma, e molte, molte altre ..

Apprendimento automatico (Machine Learning)

Branca dell' IA. Capacità di un computer di imparare a compiere una specifica azione senza essere stato programmato con specifiche istruzioni.

Classificazione: decidere se un'e-mail è indesiderata, diagnostica per immagini, ...

Regressione: profilo utenti per tracciamento pubblicitario partite in giochi specifici, guida autonoma, ...



Nel caso in cui si chieda al calcolatore di eseguire compiti semplici, si può creare un algoritmo che lo guidi all'esecuzione. Spieghiamo esattamente al calcolatore cosa fare dividendo un problema in sotto-problemi semplici e diamo istruzioni dettagliate su come risolverli.

Nel caso di compiti più complessi invece, risulta più efficiente far ricavare al calcolatore il proprio algoritmo.

Nel caso in cui si chieda al calcolatore di eseguire compiti semplici, si può creare un algoritmo che lo guidi all'esecuzione. Spieghiamo esattamente al calcolatore cosa fare dividendo un problema in sotto-problemi semplici e diamo istruzioni dettagliate su come risolverli.

Nel caso di compiti più complessi invece, risulta più efficiente far ricavare al calcolatore il proprio algoritmo.

Supervised learning Il calcolatore impara da un insieme di dati (training set) al fine di eseguire una certo compito.

Supervised learning Il calcolatore impara da un insieme di dati (training set) per eseguire una certo compito.

L'insieme di dati contiene una serie di dati di input e output corrispondenti, tipicamente organizzati in una matrice. Il calcolatore apprende da tali dati ed individua il valore di output più adatto per nuovi dati di input.

Si raggiunge tale scopo attraverso la ricerca del minimo di una opportuna funzione, detta funzione di costo (cost function).

Regressione

Idea: a partire da un insieme di dati cerchiamo un modello adatto alla rappresentazione degli stessi attraverso l'uso di parametri opportuni.

Spesso la regressione si può ripensare come la soluzione di un opportuno sistema lineare:

$$Ax=b$$

Table 1

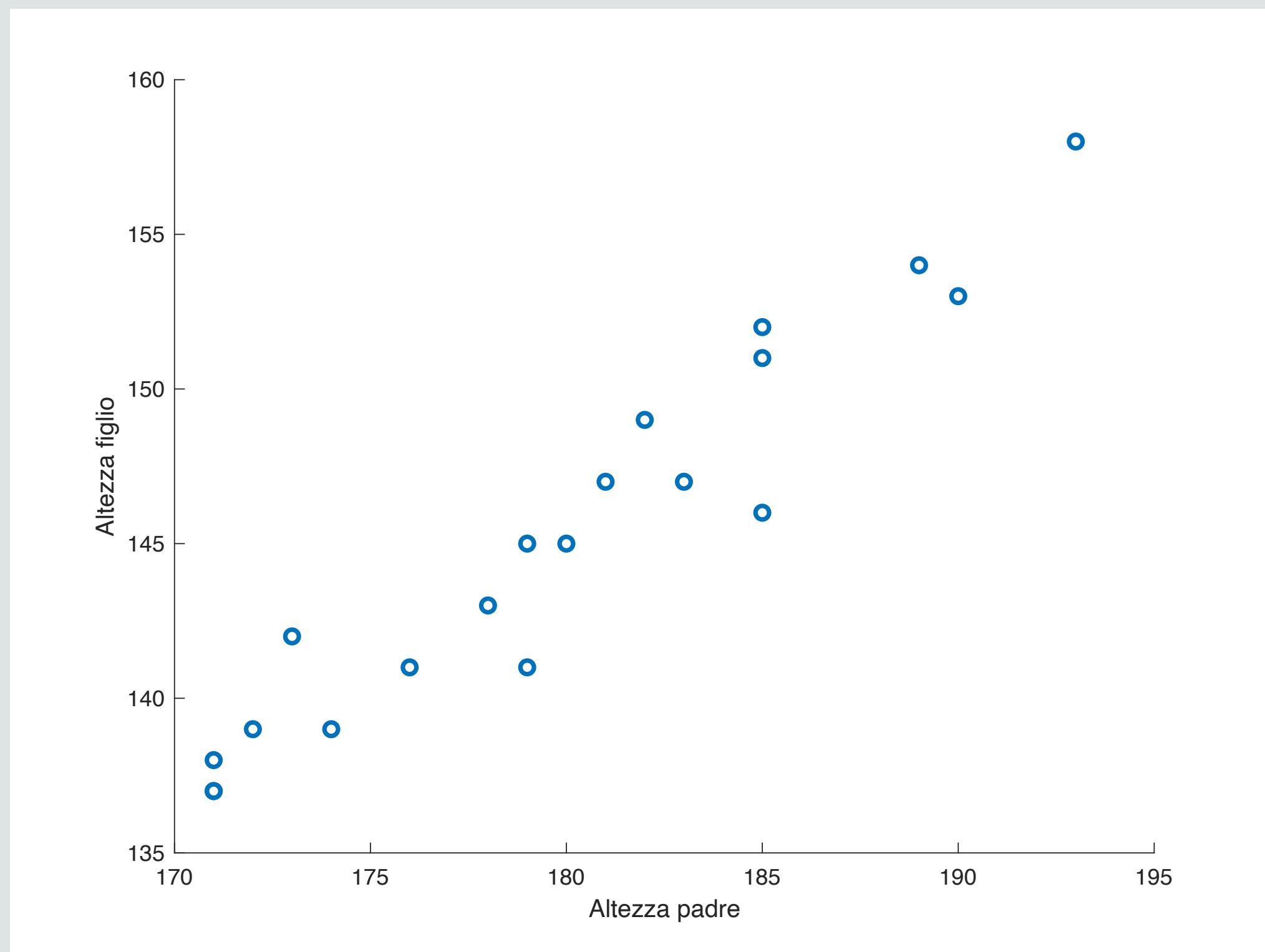
Altezza padre	Altezza studente
190	153
173	142
176	141
193	158
171	138
182	149
171	137
179	141
174	139
185	152
180	145
172	139
171	137
183	147
185	151
179	145
178	143
181	147
185	146
189	154

Esempio

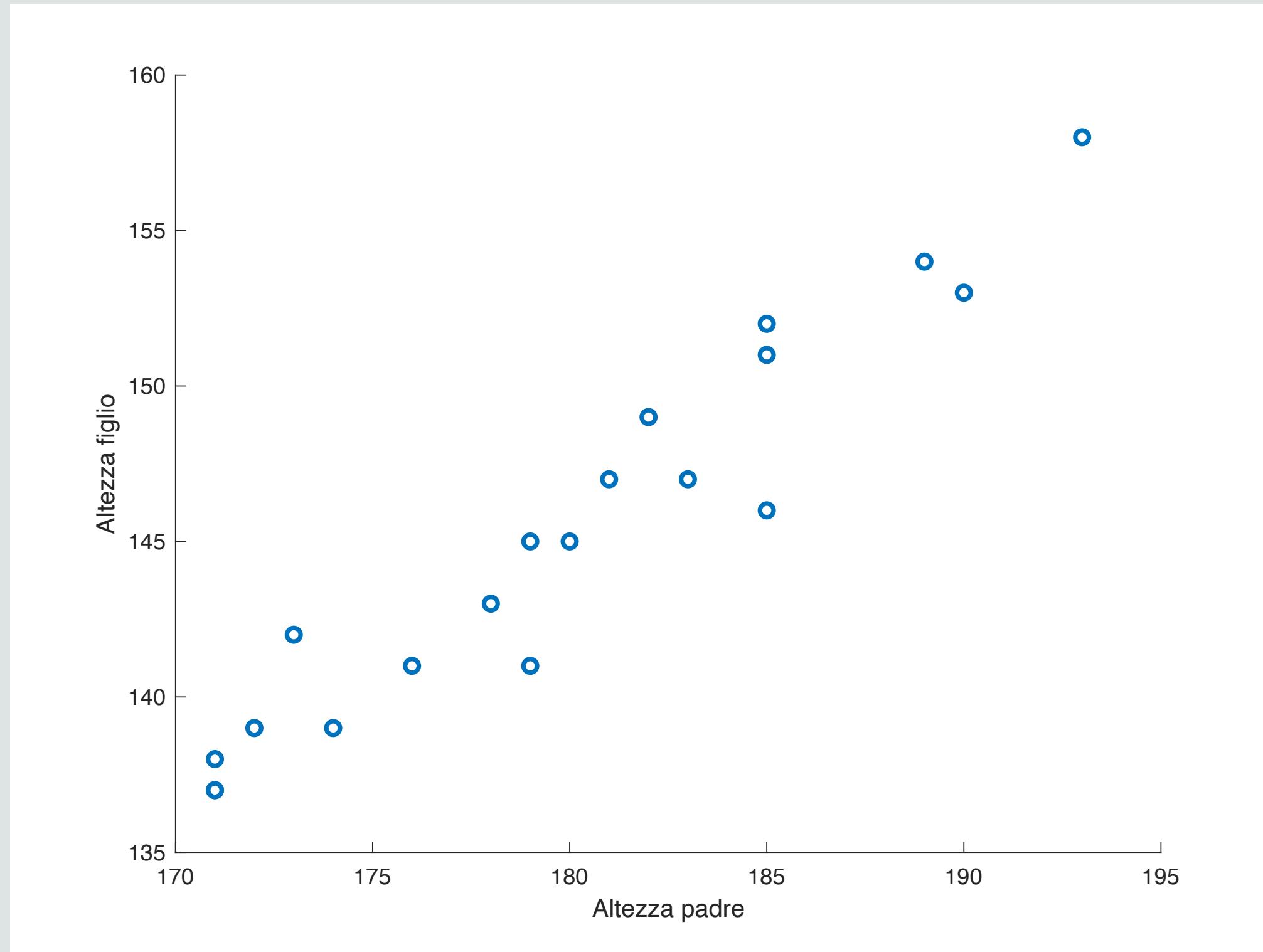
In che misura conoscere l'altezza del padre ci può aiutare a **predire** quella del figlio?

In che misura le due variabili sono **correlate**?

Table 1

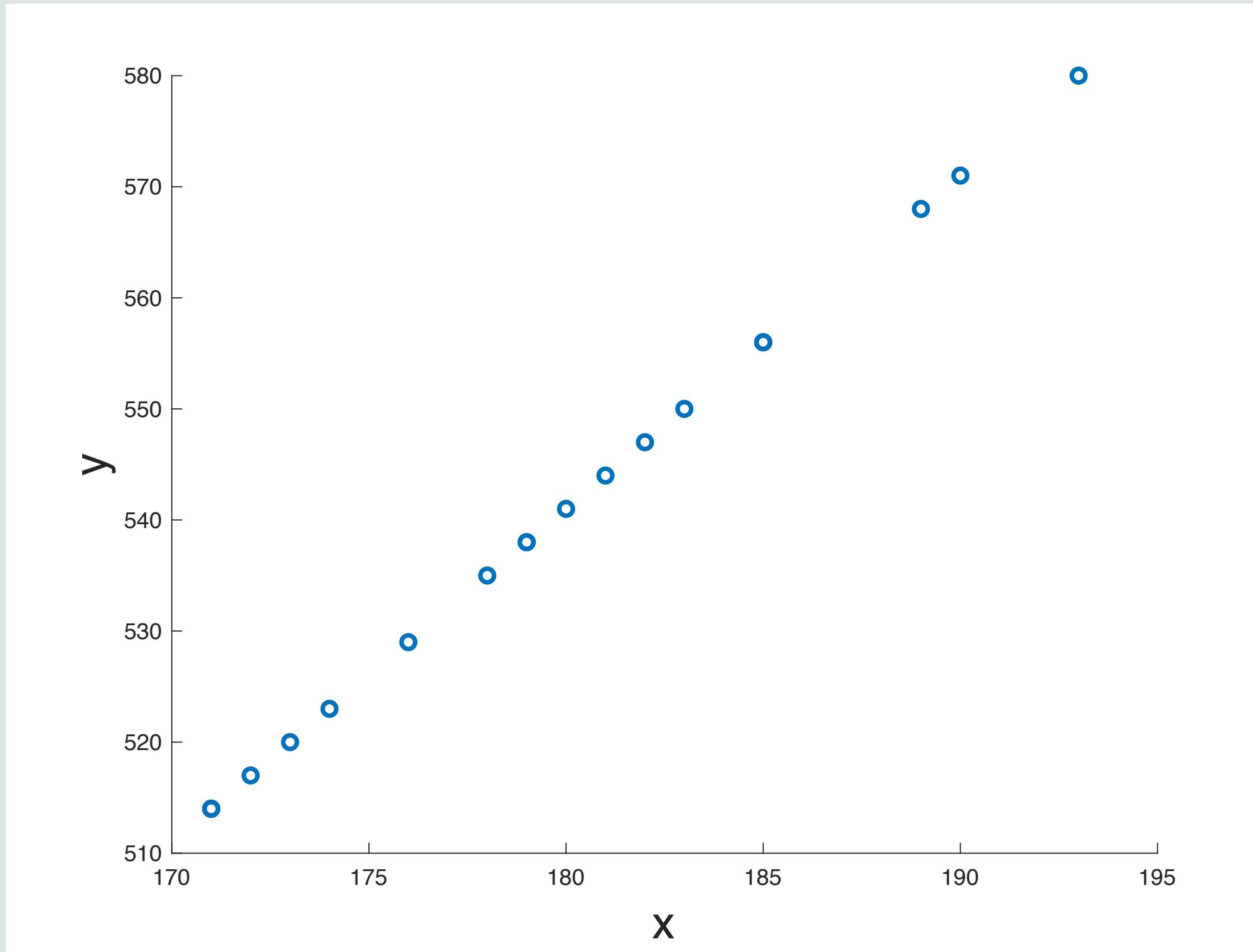


Altezza padre	Altezza studente
190	153
173	142
176	141
193	158
171	138
182	149
171	137
179	141
174	139
185	152
180	145
172	139
171	137
183	147
185	151
179	145
178	143
181	147
185	146
189	154



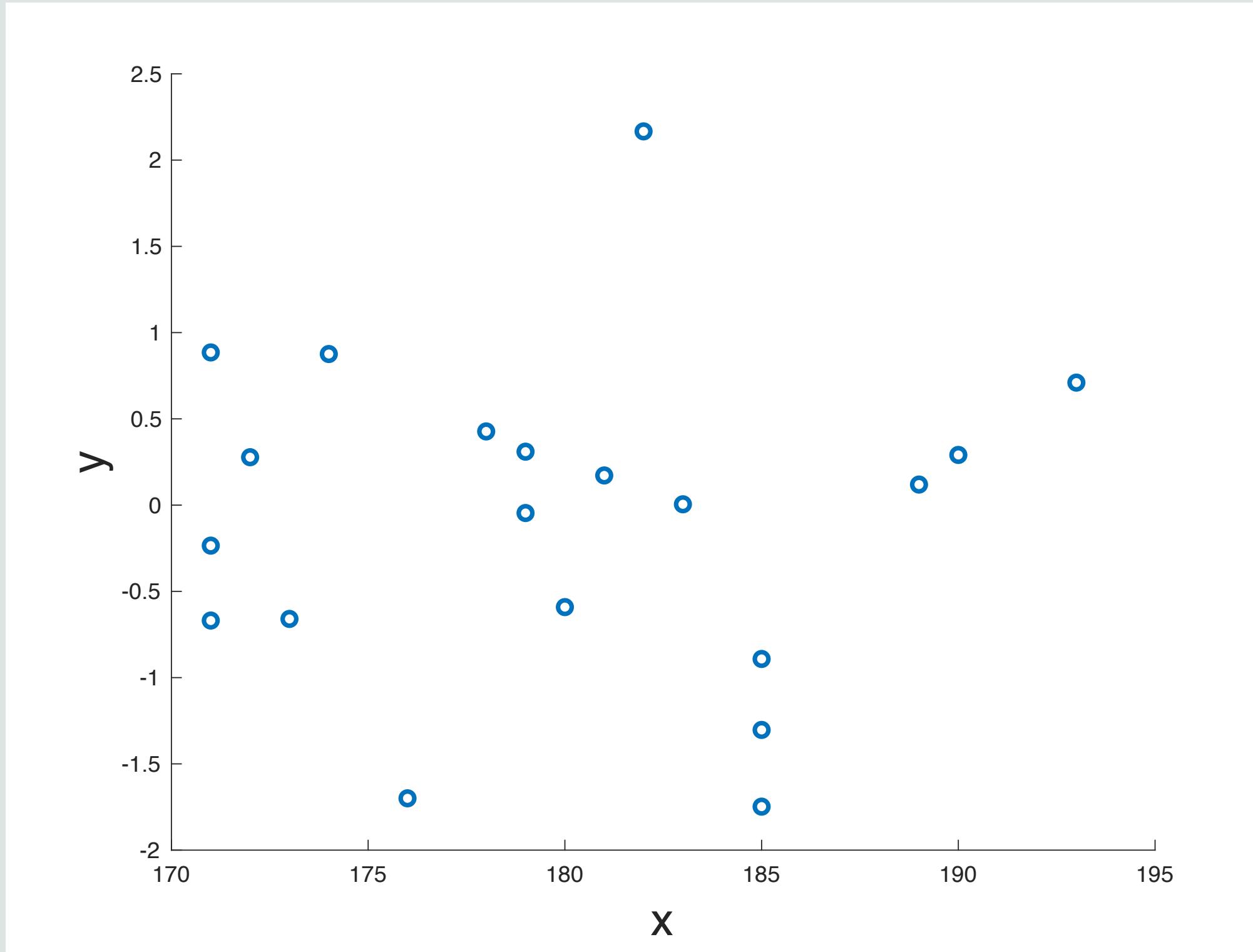
Dalla rappresentazione grafica dei dati sembra esserci una relazione lineare tra le due altezze. x ed y sembrano correlate.

Altri esempi



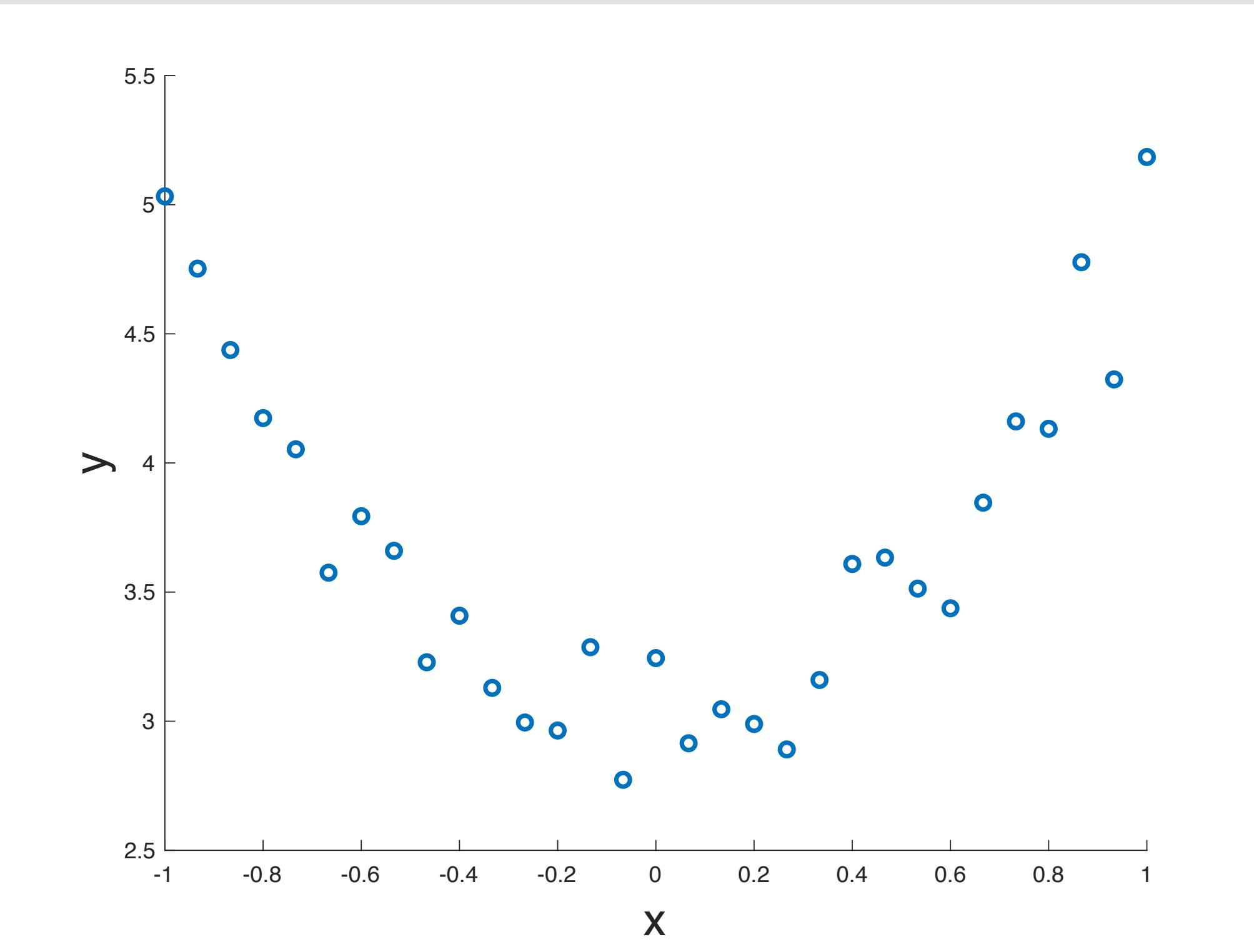
x ed y in questo
grafico sono
perfettamente
correlate

Altri esempi

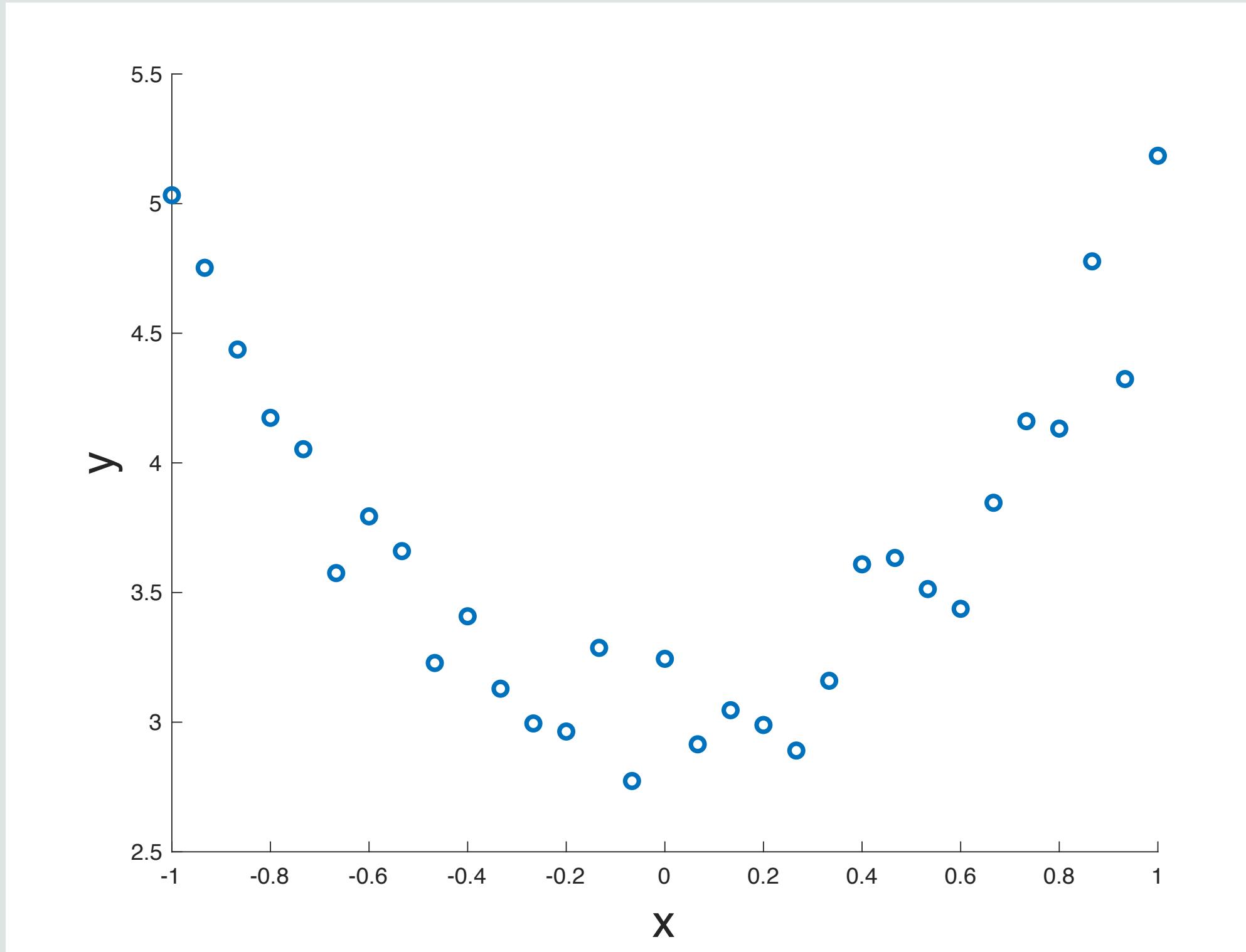


x ed y in questo
grafico non sono
correlate

Altri esempi

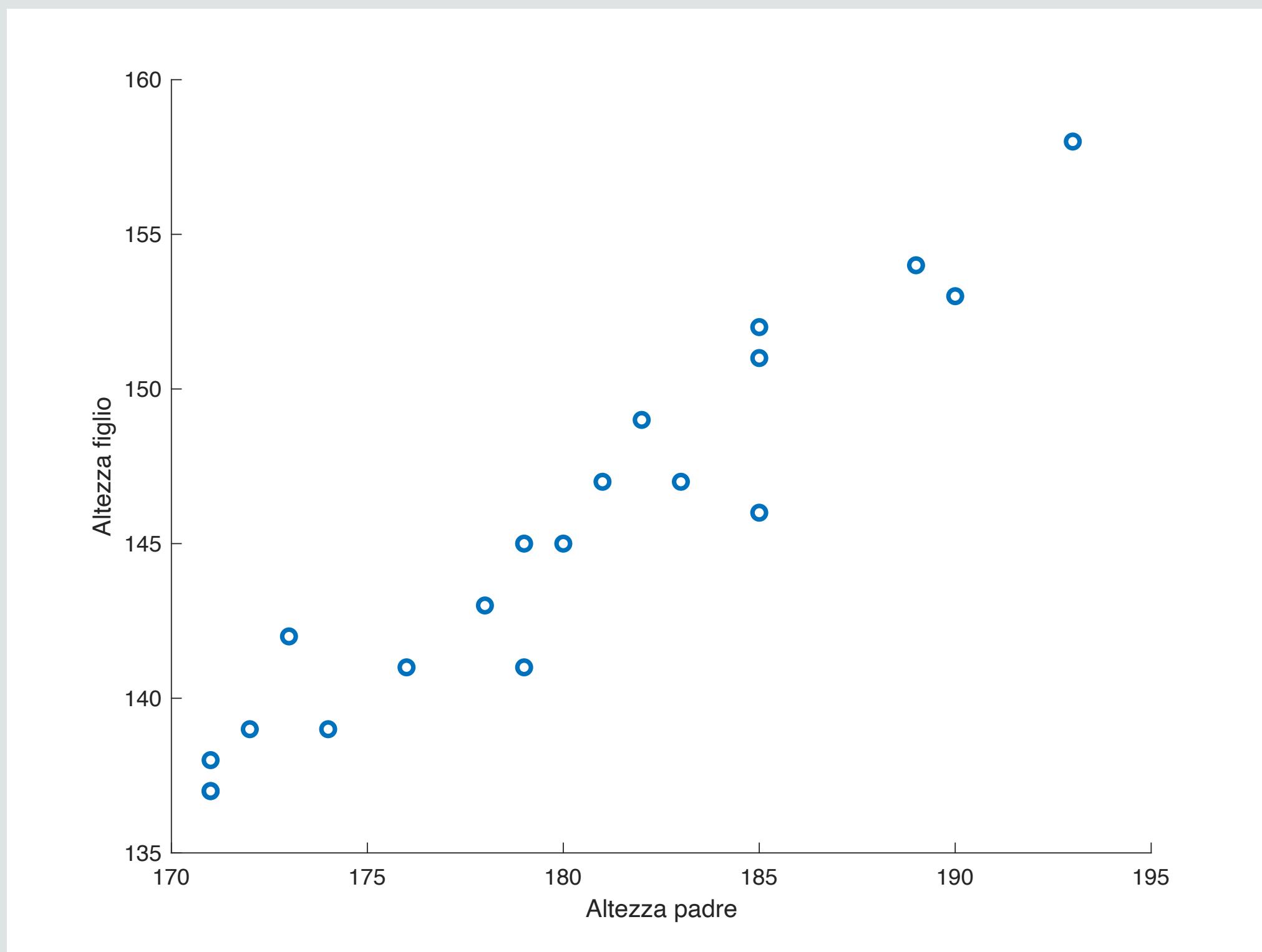


Altri esempi



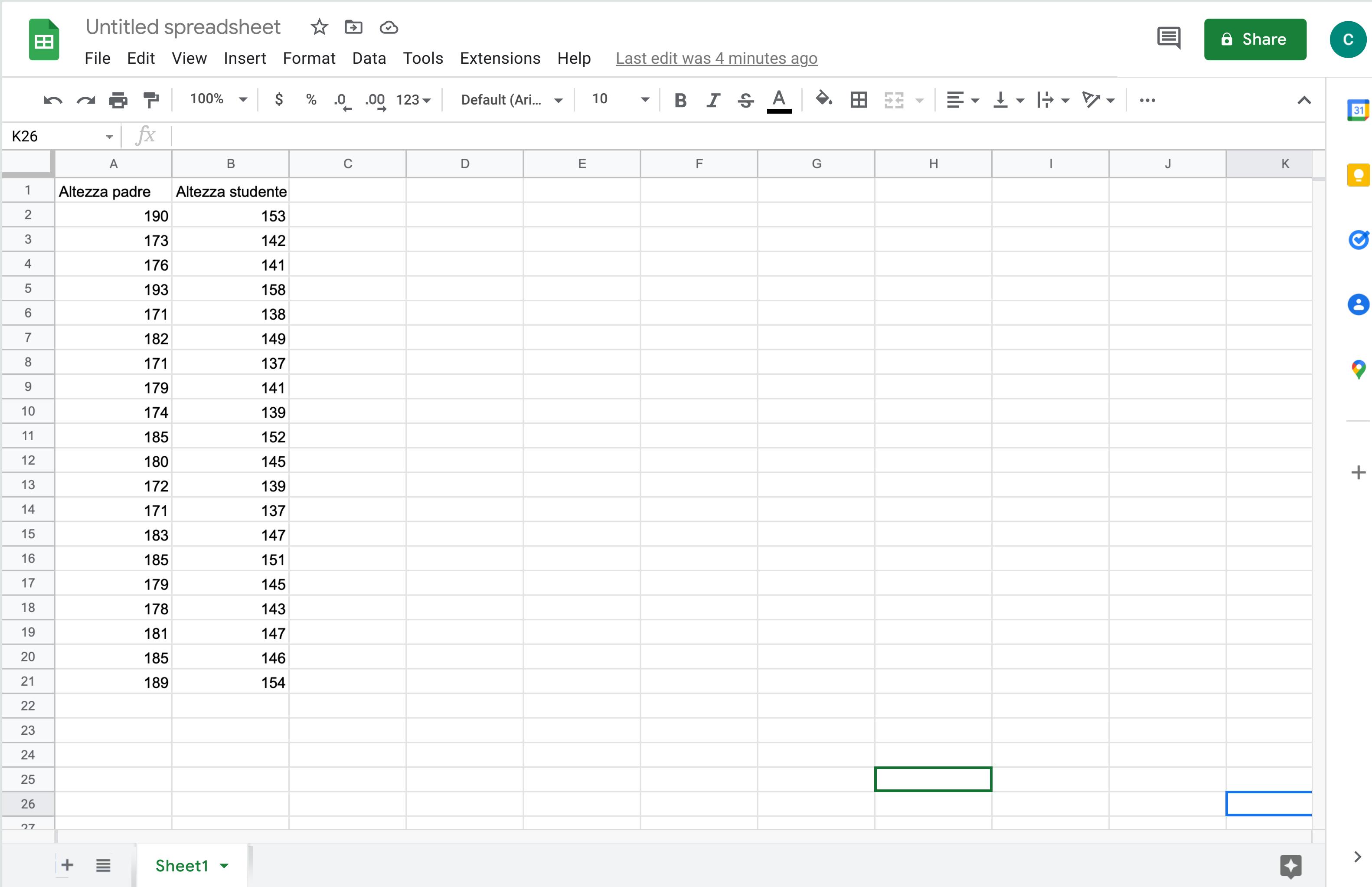
Tra x ed y in questo grafico non c'è una relazione lineare ma quadratica. Es: lato di un quadrato e area del quadrato

Table 1



Altezza padre	Altezza studente
190	153
173	142
176	141
193	158
171	138
182	149
171	137
179	141
174	139
185	152
180	145
172	139
171	137
183	147
185	151
179	145
178	143
181	147
185	146
189	154

Risolviamo il problema di regressione con l'ausilio di Google Sheets.

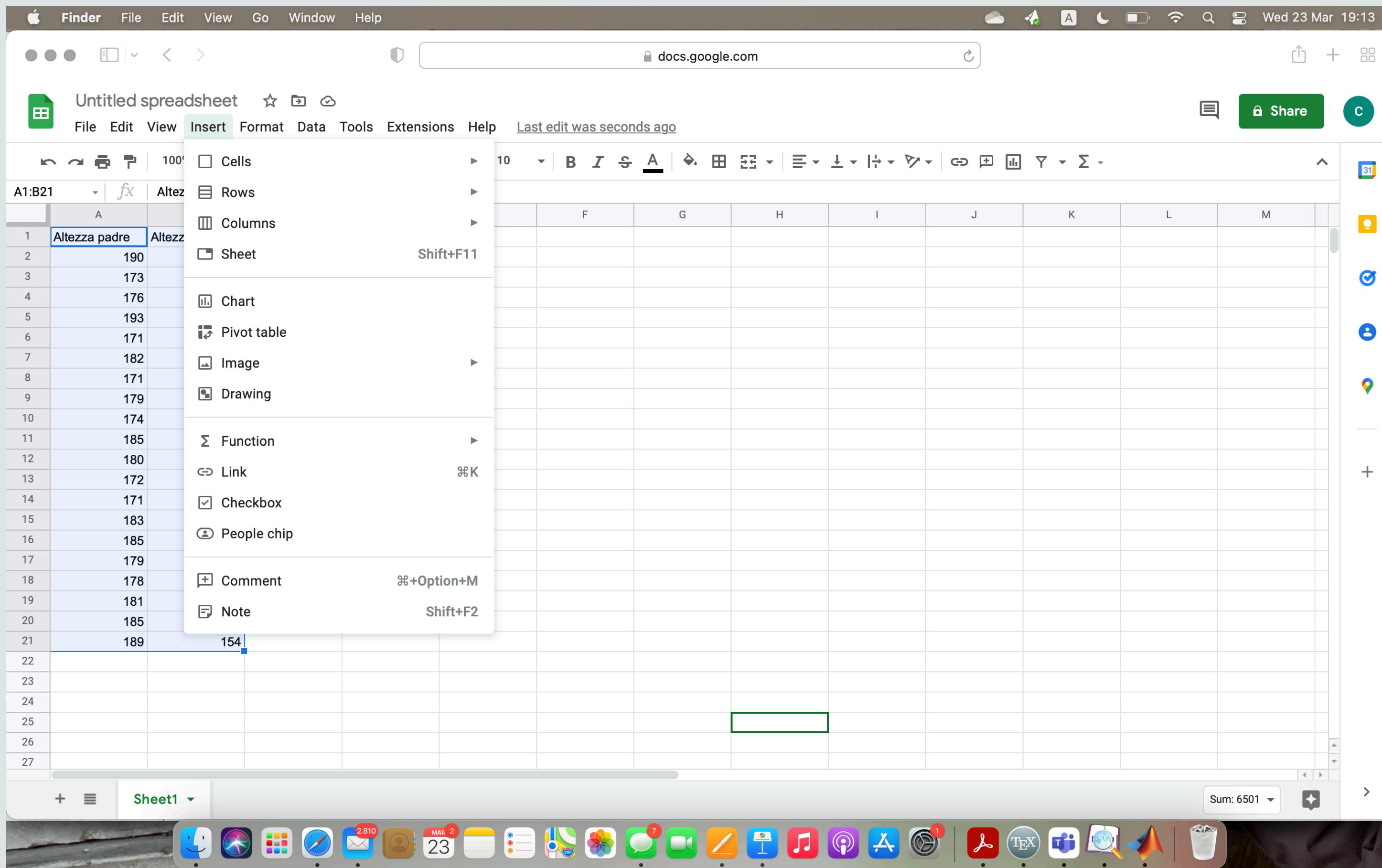


The screenshot shows a Google Sheets spreadsheet titled "Untitled spreadsheet". The data is organized into two columns: "Altezza padre" (Height of parent) in column A and "Altezza studente" (Height of student) in column B. The data starts from row 1 and continues down to row 27. The first row contains the column headers. The data shows a positive correlation, with heights ranging from 137 to 193 for parents and 137 to 154 for students. The spreadsheet interface includes a toolbar at the top with various icons for file operations, a "Share" button, and a sidebar on the right with icons for calendar, notes, and other Google services.

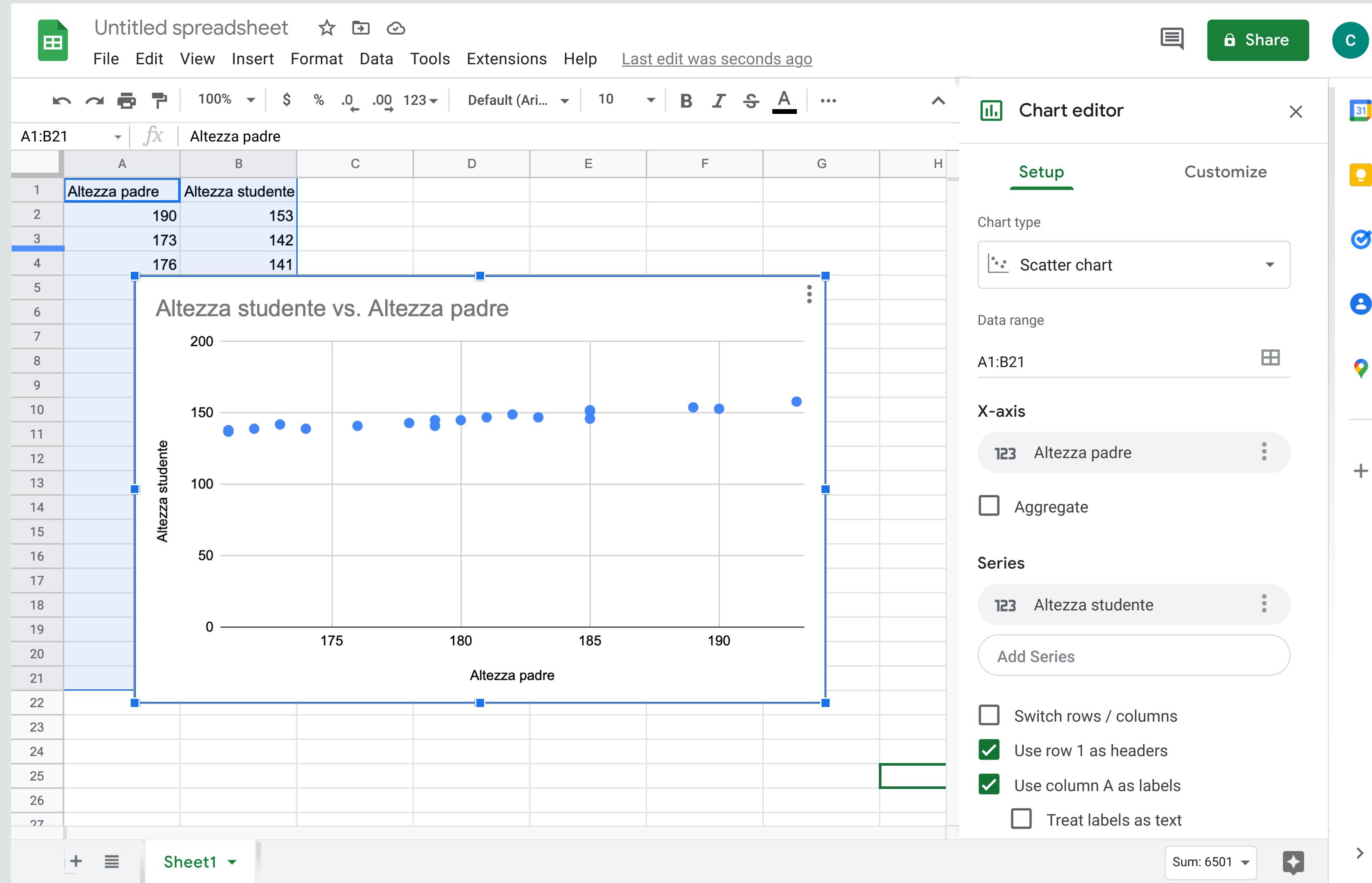
	A	B
1	Altezza padre	Altezza studente
2	190	153
3	173	142
4	176	141
5	193	158
6	171	138
7	182	149
8	171	137
9	179	141
10	174	139
11	185	152
12	180	145
13	172	139
14	171	137
15	183	147
16	185	151
17	179	145
18	178	143
19	181	147
20	185	146
21	189	154
22		
23		
24		
25		
26		
27		

Selezione i dati dalla casella A1 alla B21

Seleziono Insert e Chart dal Menù

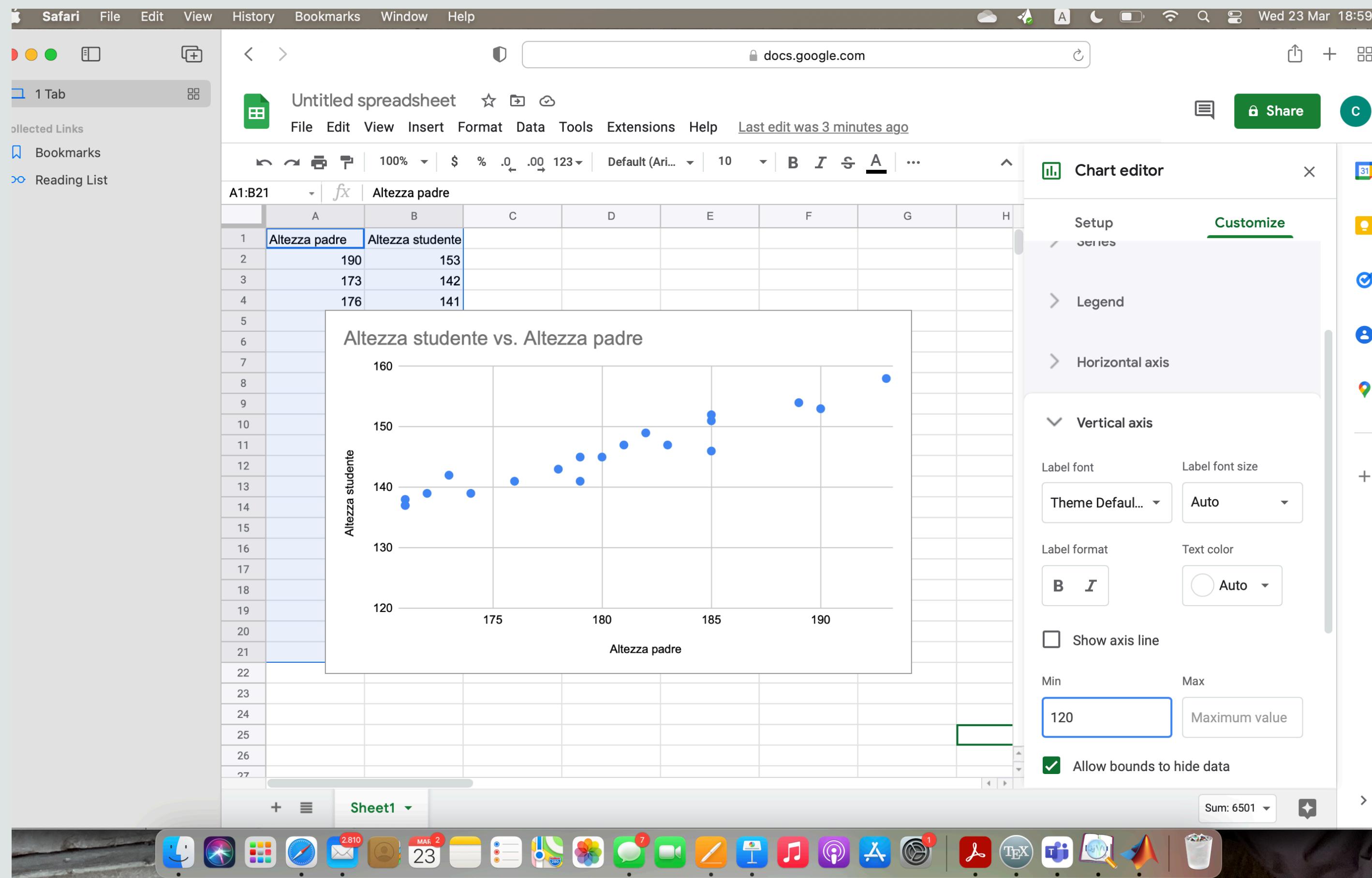


Selezione Scatter Plot

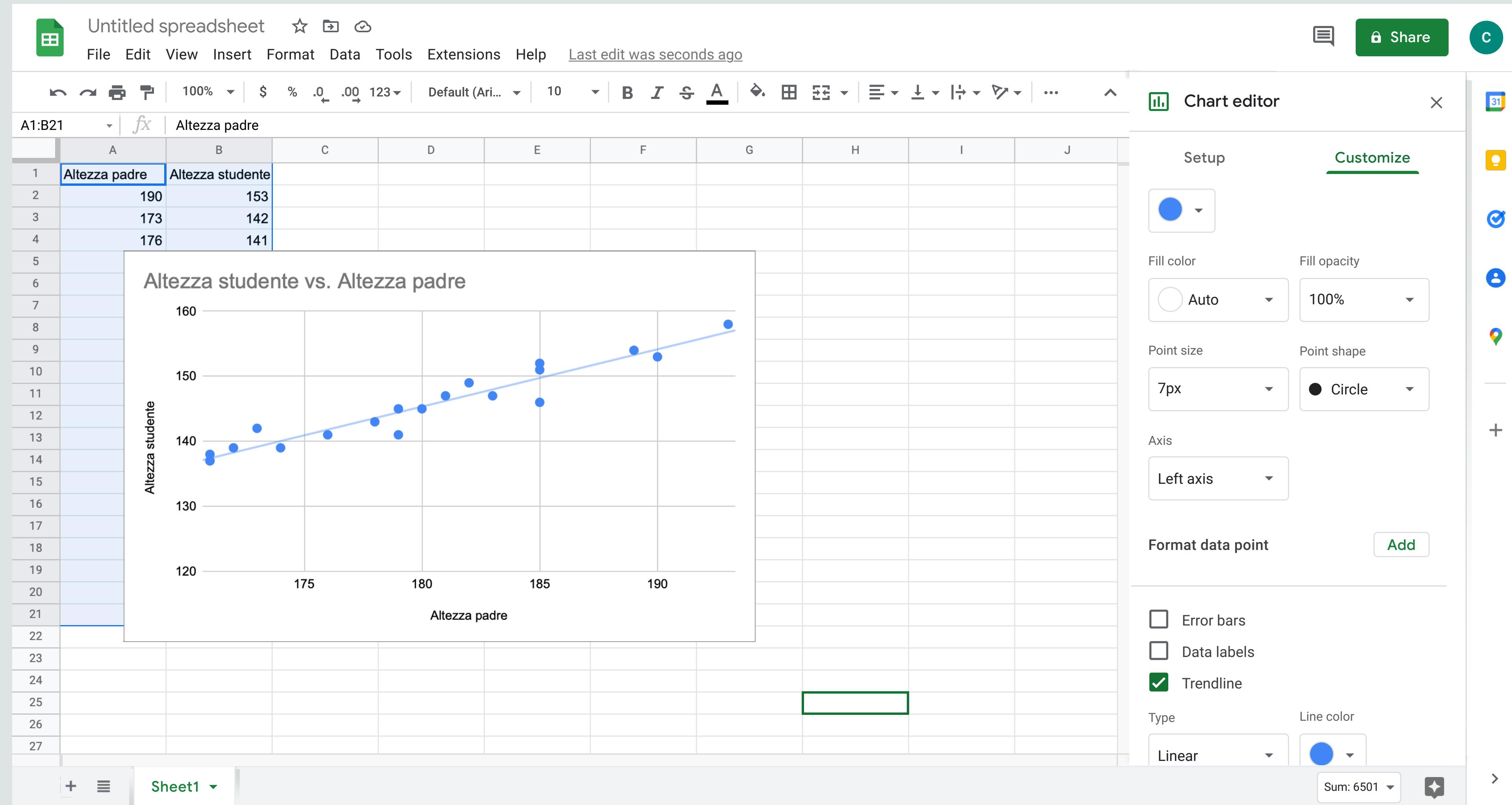


Gli estremi dell'asse verticale sono troppo lontani

Selezione Customize, Vertical axis e scelgo Min 120



Selezione Trendline e scelgo Linear



Seleziono Label e Use Equation per visualizzare l'equazione della retta di regressione lineare

