

Multiple linear regression has been around for over a century, with its roots going back to the late 19th and early 20th centuries. Here's a brief history of its development:

Early Origins (1800s):

- **Carl Friedrich Gauss (1777–1855):** Gauss was instrumental in the development of the **method of least squares**, a fundamental concept in linear regression. While his work primarily focused on simple linear regression, the foundation he laid would later be extended to multiple variables.
- **Adrien-Marie Legendre (1752–1833):** Independently of Gauss, Legendre also developed the method of least squares and published it in 1805. This method became central to estimating the parameters of linear regression models.

Birth of Multiple Regression (Late 1800s to Early 1900s):

- **Sir Francis Galton (1822–1911):** Galton, a British polymath, worked on studying the relationship between variables in biological and social sciences. His work on **regression to the mean** (around the 1880s) laid the groundwork for linear regression by exploring the relationship between a dependent variable and one or more independent variables.
- **Karl Pearson (1857–1936):** A protégé of Galton, Pearson extended regression methods and is credited with formalizing **multiple correlation** and **multiple linear regression** in the late 19th and early 20th centuries. His work in **biometry** (applying statistics to biology) was key in the development of the statistical tools used in multiple linear regression.

Key Developments in the 20th Century:

- **Ronald A. Fisher (1890–1962):** Fisher made significant contributions to the field of statistics in the early 20th century, particularly in the area of **regression analysis**. He developed many foundational tools, including the **analysis of variance (ANOVA)**, and integrated these with regression models to help explain the relationships between variables in a more sophisticated way. Fisher's contributions helped cement multiple regression as a key statistical tool.
- **1930s–1950s:** By the mid-20th century, the method of **multiple linear regression** had become a well-established statistical tool, used across many fields like economics, biology, psychology, and the social sciences. The advent of computers in the 1950s further accelerated the use of regression methods because they allowed for faster and more complex calculations.

Multiple Linear Regression in the Modern Era:

- In the **latter half of the 20th century**, multiple linear regression became a standard method for analyzing relationships between variables in various disciplines, including economics, engineering, medicine, and the social sciences.
- Today, **multiple linear regression** is a foundational statistical method taught in introductory statistics courses and is used in everything from **predictive modeling** to **machine learning** and **data science**.

Summary of Timeline:

- **1805:** The **method of least squares** (underpinning linear regression) was independently developed by Carl Friedrich Gauss and Adrien-Marie Legendre.
- **Late 1800s:** Sir Francis Galton introduced the idea of regression to the mean.
- **Early 1900s:** Karl Pearson formalized multiple correlation and multiple linear regression.
- **1920s–1930s:** Ronald A. Fisher expanded regression methods to incorporate more complex analyses, such as ANOVA.
- **1950s onward:** With the rise of computers, multiple linear regression became a widely used and essential tool in applied statistics and data analysis.

Multiple linear regression has thus been around for over **100 years**, evolving from early work on least squares estimation to become one of the most commonly used statistical tools for understanding relationships between variables in a wide range of fields.