Sum of Squares Total, Sum of Squares Regression and Sum of Squares Error

Statistics Tutorials 5 min read

You may be wondering what all of those **sums of squares** are all about. Maybe that's what got you here in the first place. Well, they are the determinants of a good linear **regression** (https://365datascience.com/linear-regression/). This tutorial is based on the ANOVA framework (https://www.xlstat.com/en/solutions/features/anova-analysis-of-variance) you may have heard before.

Before reading it, though, make sure you are not mistaking **regression** for **correlation** (https://365datascience.com/correlation-regression/). If you've got this checked, we can get straight into the action.

A quick side-note: Want to learn more about linear regression? Check out our explainer videos The Linear Regression Model. Geometrical Representation

Related Posts

Exploring the OLS Assumptions (https://365datascience.com/ols assumptions/)

Introducing the Normal Distribution (https://365datascience.com/nor distribution/)

How to Visualize Numerical Data with Histograms (https://365datascience.com/nur data-histogram/)

What is the Student's T Distribution? (https://365datascience.com/stu t-distribution/)

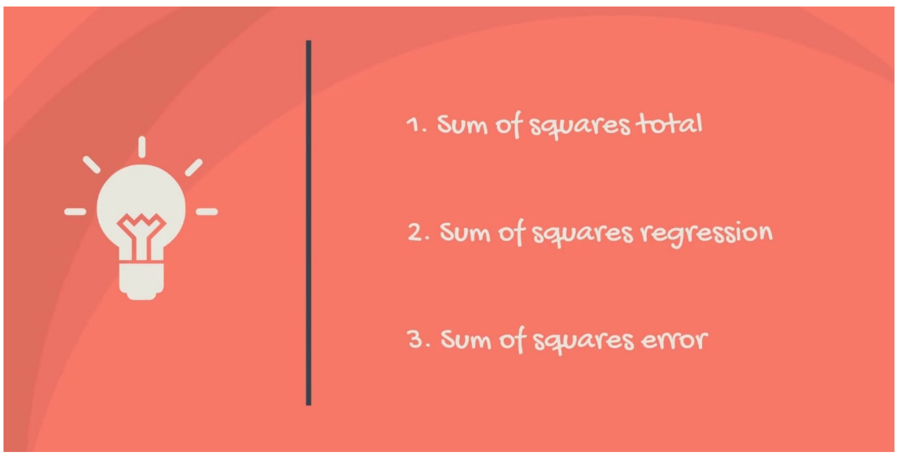
Introduction to the Measures of Central Tendency

(https://365datascience.com/explainer-video/linear-regression-model/) and The Simple Linear Regression Model (https://365datascience.com/explainer-video/simple-linear-regression-model/).

(https://365datascience.com/me central-tendency/)

SST, SSR, SSE: Definition and Formulas

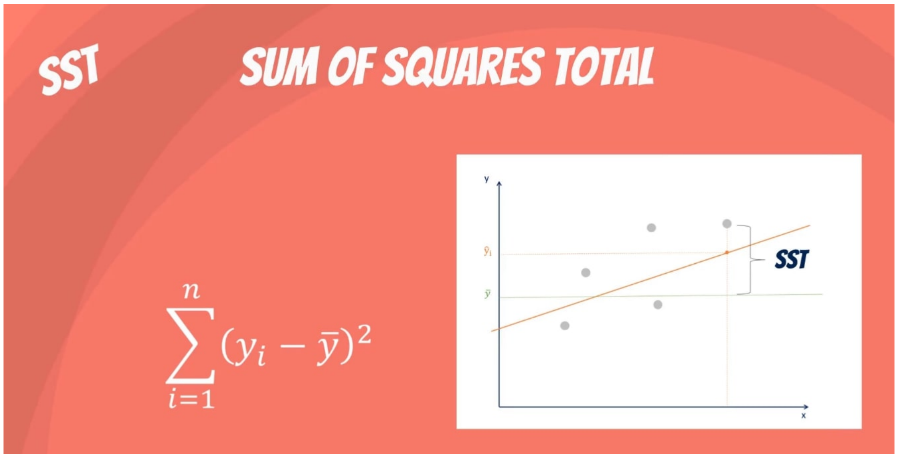
There are three terms we must define. The **sum of squares total**, the **sum of squares regression**, and the **sum of squares error**.



(https://365datascience.com/wp-content/uploads/2018/11/image1-6.jpg)

What is the SST?

The **sum of squares total**, denoted **SST**, is the squared differences between the observed *dependent variable* and its **mean**. You can think of this as the dispersion of the observed variables around the **mean** (https://365datascience.com/measures-central-tendency/) – much like the **variance** in descriptive statistics (https://365datascience.com/coefficient-variation-variance-standard-deviation/).



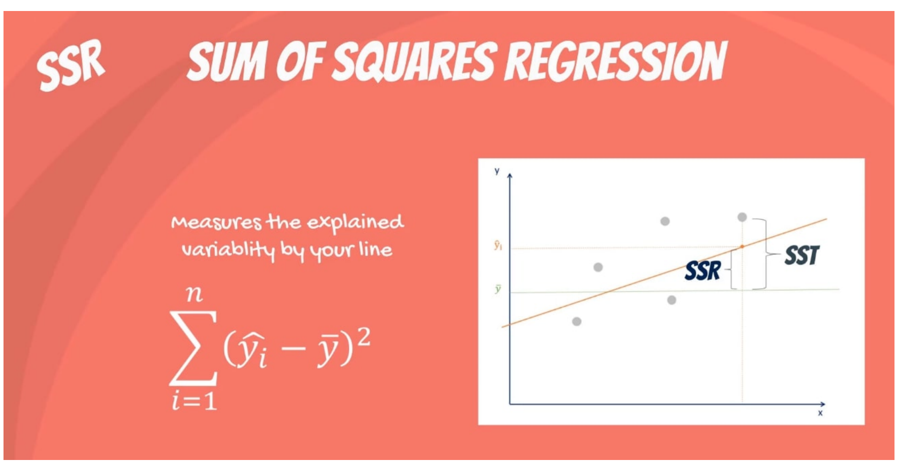
(https://365datascience.com/wp-content/uploads/2018/11/image2-6.jpg)

It is a measure of the total variability of the dataset.

Side note: There is another notation for the **SST**. It is **TSS** or **total sum of squares**.

What is the SSR?

The second term is the **sum of squares due to regression**, or **SSR**. It is the sum of the differences between the *predicted* value and the **mean** of the *dependent variable*. Think of it as a measure that describes how well our line fits the data (https://365datascience.com/numerical-categorical-data/).

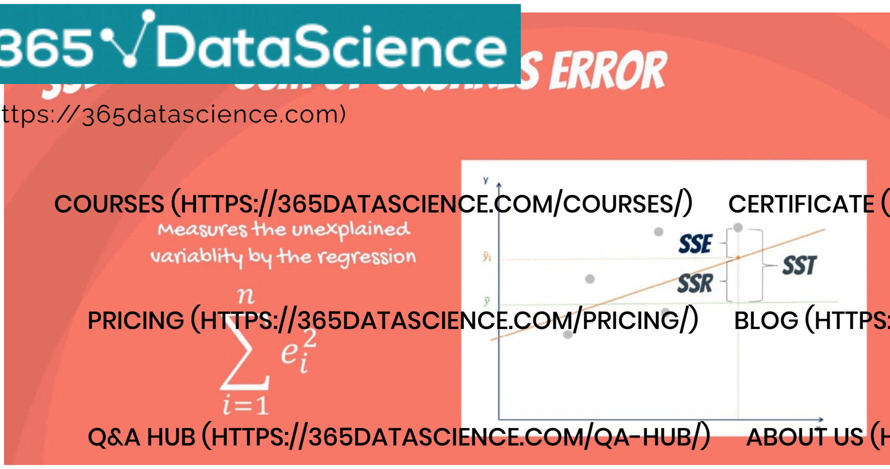


(https://365datascience.com/wp-content/uploads/2018/11/image3-6.jpg)

If this value of **SSR** is equal to the **sum of squares total**, it means our **regression model** captures all the observed variability and is perfect. Once again, we have to mention that another common notation is **ESS** or **explained sum of squares**.

What is the SSE?

The last term is the **sum of squares error**, or **SSE**. The error is the difference between the *observed* value and the *predicted* value.



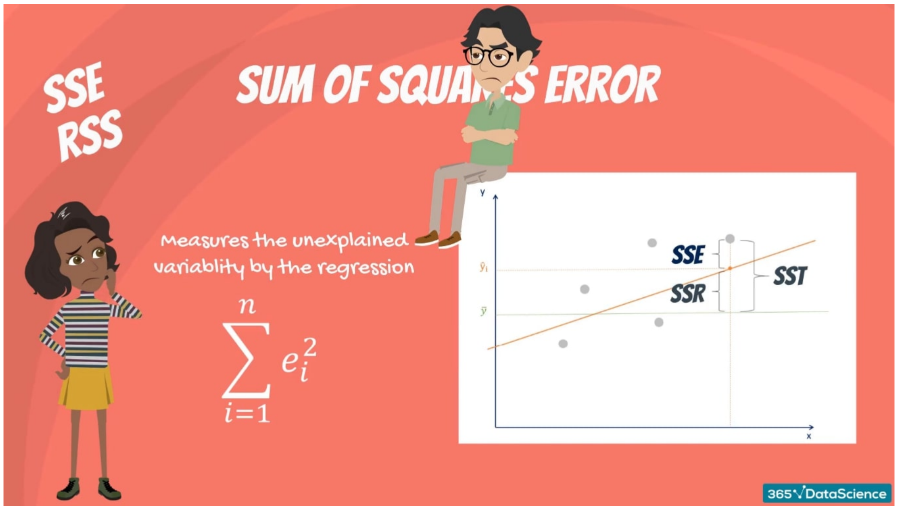
(https://365datascience.com/wp-content/uploads/2018/11/image4-6.jpg)

We usually want to minimize the error (https://365datascience.com/false-positive-vs-false-negative/). The smaller the error, the better the estimation power of the **regression**. Finally, I should add that it is also known as **RSS** or **residual sum of squares**. Residual as in: remaining or unexplained.

The Confusion between the Different Abbreviations

It becomes really confusing because some people denote it as **SSR**. This makes it unclear whether we are talking about the **sum of squares due to regression** or **sum of squared**

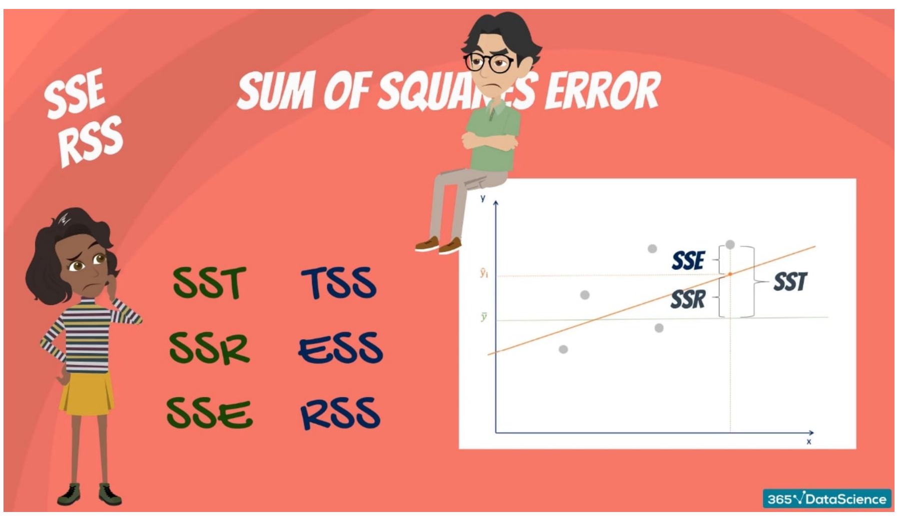
residuals.



(https://365datascience.com/wp-content/uploads/2018/11/image5-5.jpg)

In any case, neither of these are universally adopted, so the confusion remains and we'll have to live with it.

Simply remember that the two notations are **SST**, **SSR**, **SSE**, or **TSS**, **ESS**, **RSS**.

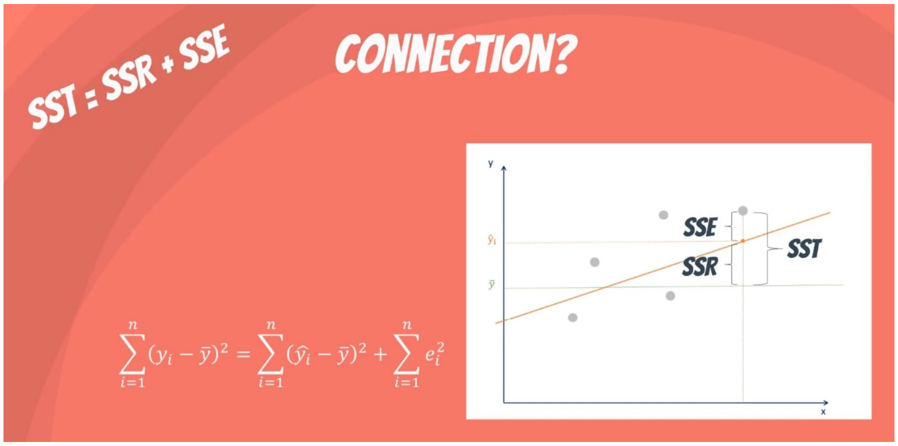


(https://365datascience.com/wp-content/uploads/2018/11/image6-5.jpg)

There's a conflict regarding the abbreviations, but not about the concept and its application. So, let's focus on that.

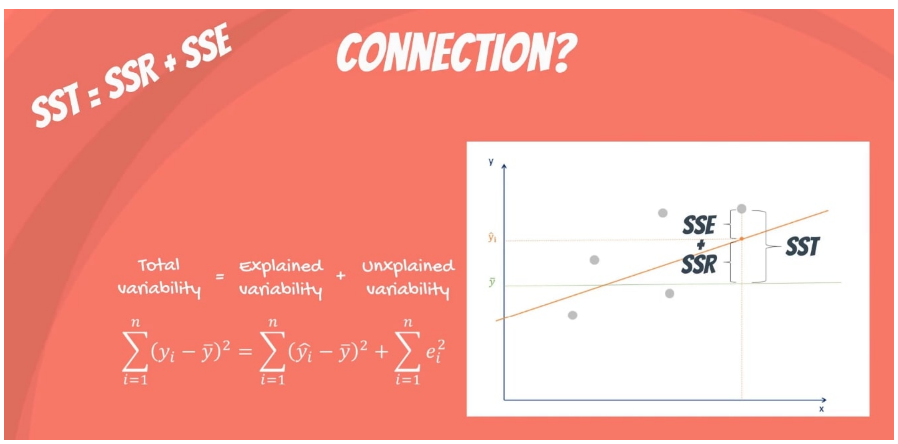
How Are They Related?

Mathematically, **SST** = **SSR** + **SSE**.



(https://365datascience.com/wp-content/uploads/2018/11/image7-5.jpg)

The rationale is the following: the total variability of the data set is equal to the variability explained by the **regression line** plus the unexplained variability, known as error.



(https://365datascience.com/wp-content/uploads/2018/11/image8-5.jpg)

Given a constant total variability, a lower error will cause a better **regression**. Conversely, a higher error will cause a less powerful **regression**. And that's what you must remember, no matter the notation.

Next Step: The R-squared

Well, if you are not sure why we need all those **sums of squares**, we have just the right tool for you. The **R-squared.** Care to learn more? Just dive into the linked tutorial where you will understand how it measures the explanatory power of a linear regression!

***

Interested in learning more? You can take your skills from good to great with our statistics course.

Try statistics course for free (https://365datascience.com/courses/statistics/)

Next Tutorial: Measuring Variability with the R-squared (https://365datascience.com/r-squared/)

Related Posts

(https://365datascience.com/point-estimates-confidence-intervals/)

How to Use Point Estimates and Confidence Intervals (https://365datascience.com/point-estimates-confidence-intervals/)

7 min rea Statistics Tutorials

(https://365datascience.com/measures-central-tendency/)

Introduction to the Measures of Central Tendency (https://365datascience.com/measures-central-tendency/)

6 min read Statistics Tutorials

(https://365datascience.com/numerical-categorical-data/)

Examples of Numerical and Categorical Variables (https://365datascience.com/numerical-categorical-data/)

5 min read Statistics Tutorials

(https://365datascience.com/measuring-asymmetry-skewness/)

Measuring Asymmetry with Skewness (https://365datascience.com/measuring-asymmetry-skewness/)

4 min read Statistics Tutorials

Leave a Reply

Your email address will not be published.

COMMENT

NAME

EMAIL

POST COMMENT

ttps://365datascience.com)

(https://www.youtube.com/channel/UCEBpSZhI1X8WaP- (https://www.linkedin.com/company-
ttps://www.facebook.com/365datascience/) kY_2LWbg) (https://twitter.com/365DataScience)

Pricing (https://365datascience.com/pricing/)

Courses (/courses/)

ources (https://365datascience.com/resources/)

olarship (/scholarship/)

lainer videos (/explainer-videos/)

nts (https://aigents.co/opportunities/events)

lainer videos (https://365datascience.com/explainer-videos/)

el (/explainer-videos/#excel)

tistics (/explainer-videos/#statistics)

gramming (/explainer-videos/#programming)

chine learning (/explainer-videos/#machine-learning)

g (https://365datascience.com/blog/)

orials (https://365datascience.com/blog/#tutorials)

eer Advice (https://365datascience.com/blog/#career-advice)

rviews (https://365datascience.com/blog/#interviews)

out us (https://365datascience.com/about-us/)

ntact us (/about-us/#contact-form)

acy Policy (https://365datascience.com/privacy-policy/)