Appendix B

Notation

This appendix contains a summary of the notation used in this book. The tables are provided for easy reference: whenever you encounter some math symbol whose meaning you don't know, you can look it up here to learn how to read it and what it means.

Math notation

Math notation is a precise language for describing variables, operations, equations, and functions. We'll now provide a condensed summary of the notation used for basic math concepts intended as a review for readers whose math skills might be a little *rusty*.

Variables and constants The usual names we use for variables are from the end of the alphabet x, y, z, m, n, etc. In contrast, we denote constants using letters from the beginning of the alphabet a, b, c, or using Greek letters. Yeah, we'll use the Greek alphabet a lot, but I don't want you to freak out about this. For example, the population mean is denoted μ (the Greek letter *mu*), and the population standard deviation is denoted σ (the Greek letter *sigma*). The Greek letters *alpha* α and *beta* β are the equivalents of a and b. Other Greek letters we'll use in this book include: *lambda* λ, *delta* Δ, *chi* χ (pronounced *khai*), *epsilon* ϵ, *gamma* γ, *nu* ν, *rho* ρ, *tau* τ, and *theta* θ.

We'll use the subscript notation to denote several variables of the same type, for example we can denote three observations of the variable x as x_1, x_2, x_3.

Math operations The book uses the standard math notation for arithmetic operations, as summarized in the following table.

Expression	Read as	Geometric interpretation
$a+b$	a plus b	sum of lengths a and b
$a-b$	a minus b	difference between lengths a and b
$-a$	negative a	same length as a in the opposite direction
$a \cdot b = ab$	a times b	area of a rectangle with sides a and b
$a^2 = aa$	a squared	area of a square of side length a
$a^3 = aaa$	a cubed	volume of a cube of side length a
a^n	a exponent n	a multiplied by itself n times
$\sqrt{a} = a^{\frac{1}{2}}$	square root of a	the side length of a square of area a
$a/b = \frac{a}{b}$	a divided by b	a parts of a length split into b parts
$a^{-1} = \frac{1}{a}$	one over a	division by a
%	percent	proportions of a total; $a\% \stackrel{\text{def}}{=} \frac{a}{100}$

We denote multiplication between the numbers a and b using the centre-dot symbol $a \cdot b$, or sometimes the multiplication symbol $a \times b$, but most commonly don't use any symbol at all. The expression ab uses *implicit multiplication*, since the default operation when two variables are placed side-by-side is that they are multiplied together.

Equations and inequalities Most mathematical statements describe some relation between two math expressions. Here are the most common types of relations we'll see in this book.

Symbol	Read as	Used to denote
$=$	is equal to	expressions that have the same value
$\stackrel{\text{def}}{=}$	is defined as	a new variable definition
$\approx$	is approximately	expressions that are almost equal
$<$	less than	$a < b$ says a is strictly less than b
$\leqslant$	less than or equal	$a \leqslant b$ says a is less than or equal to b
$>$	greater than	$a > b$ says a is strictly greater than b
$\geqslant$	greater than or equal	$a \geqslant b$ says a is greater than or equal to b

An equation is a mathematical statement that says the expression on the left of the equality sign "$=$" is equal to the expression on the right side of the equality sign. The symbol $\stackrel{\text{def}}{=}$, read "equal by definition," is a special kind of equal sign that we use to define a new variable. For example, the statement $\overline{\mathbf{x}} \stackrel{\text{def}}{=} \mathbf{Mean}(\mathbf{x})$ defines the new symbol $\overline{\mathbf{x}}$, which is the value of the mean of the sample $\mathbf{x}$.

We use the symbol $\approx$ to denote approximations, which are similar to equations, but the equality doesn't hold exactly—the left side and the right side are only approximately equal. For example, the fraction $\frac{5}{6}$ corresponds to an infinitely long decimal $\frac{5}{6} = 0.83333333333\ldots$. In order for the equality to hold, we would need to write infinitely many 3s in the decimal expansion, which is impractical. Instead, we can write $\frac{5}{6} \approx 0.833$, which tells us the two quantities are approximately equal.

Sets A *set* is a collection of math objects. We denote sets using the curly brackets {`<desc>`}, where `<desc>` describes the elements in this set. For example the expression $S \stackrel{\text{def}}{=} \{1,2,3\}$ defines the set S that consists of the three numbers 1, 2, and 3. We can also describe sets using *interval notation*, which specifies a range of numbers using its endpoints enclosed in [or (brackets. For example, the interval $[a,b]$ describes the set of numbers between a and b, including the endpoints a and b. The interval $[a,b)$ describes the set of numbers between a and b that includes the endpoint a but not the endpoint b.

Here are some important number sets for which mathematicians have developed a special notation with double-line symbols.

Symbol	Name	Used to denote
$\mathbb{N}$	naturals	the set of natural numbers $\mathbb{N} \stackrel{\text{def}}{=} \{0,1,2,3,\ldots\}$
$\mathbb{Z}$	integers	the set of integers $\mathbb{Z} \stackrel{\text{def}}{=} \{\ldots,-2,-1,0,1,2,3,\ldots\}$
$\mathbb{Q}$	rationals	the set of fractions $\frac{m}{n}$ where m and n are integers
$\mathbb{R}$	reals	the set of real numbers $\mathbb{R} \stackrel{\text{def}}{=} (-\infty,\infty)$
$\mathbb{R}_+$	non-neg. reals	the set of non-negative real numbers $\mathbb{R}_+ \stackrel{\text{def}}{=} [0,\infty)$

All the number sets listed above contain an infinite number of elements. The notation $\ldots$ in the definition of the naturals describes an infinite sequence of integers. The set of real numbers is the most general set of numbers that includes all the naturals $\mathbb{N}$, the integers $\mathbb{Z}$, the rationals $\mathbb{Q}$, as well as irrational numbers like π, e, $\sqrt{2}$, etc. Essentially, the set of real numbers includes any numbers you might encounter in the real world.

There are two types of number objects in `Python`: integers (`int`) and floating point numbers (`float`). The the set of integers $\mathbb{Z} = \{\ldots,-2,-1,0,1,2,\ldots\}$ can are represented as Python `int` objects, while the reals $\mathbb{R}$ can be approximately represented as `float` objects.

Sets notation One of the most intimidating aspect of math notation are the symbols mathematicians use to describe set relations ($\in$, $\notin$, $\subset$) and set operations ($\cup$, $\cap$, $\setminus$). One of my students referred to these symbols as the "alien notation," and this is an accurate description if you're seeing this stuff for the first time.

I've tried to limit the use of these alien symbols to a minimum in the main text, but I use set notation in some of the problem solutions, so I figured I should include here a condensed summary that explains how to read each symbol and what it means.

Symbol	Read as	Meaning
$a \in S$	a in S	a is an element of the set S
$a \notin S$	a not in S	a is not an element of the set S
$\subset$	subset	one set strictly contained in another
$\subseteq$	subset or equal	containment or equality
$\cup$	union	the combined elements from two sets
$\cap$	intersection	the elements two sets have in common
$S \backslash T$	S set minus T	the elements of S that are not in T
$\forall x$	for all x	a statement that holds for all x
$\exists x$	there exists x	an existence statement
$\nexists x$	there doesn't exist x	a non-existence statement
$\mid$	such that	describe or restrict the elements of a set

The purpose of these alien symbols is to express mathematical statements as precisely and as concisely as possible. Whenever you want to say something using mathematics, it's important to make your statements as concise as possible, so they can fit on a single line. Short, precise statements make it easy to "see" what is going on. This is why mathematicians came up with all those symbols: they are not trying to make life difficult for you, on the contrary, they are trying want to make things as simple as possible.

This condensed symbolic notation is most useful in proofs and derivations, which are sequences of math statement that show why some statement is true by reducing or simplifying some expression. It would be very annoying to have to write every mathematical proof in natural language, writing out mathematical statement in as English sentences. Math proofs would take entire pages! Basically, you have to trust me that these complicated math symbols actually make things simpler, not more complicated.

Let's look at some examples of the alien symbols in use. The symbols $\in$ (read *in*) and $\notin$ (read *not in*) are used to describe set membership. For example, the statement $\sqrt{2} \in \mathbb{R}$ reads "the square root of two is an element of the set of real numbers," or "the square root of two is a real number." The statement $\sqrt{2} \notin \mathbb{Q}$ reads "the square root of two is not an element of the set of rational numbers," or more concisely, "the square root of two is not a rational number."

The subset symbols allow us to concisely describe which set containment. For example, the math statement $\mathbb{Q} \subset \mathbb{R}$ reads "the set of rational numbers is a subset of the set of real numbers," or more concisely "the reals contain the rationals." This means every rational number $q \in \mathbb{Q}$ is also a rational number $q \in \mathbb{R}$.

Set builder notation The *set-builder* notation is a common pattern mathematicians use to describe subsets. We can describe an arbitrary subset of the set S using the notation $\{ x \in S \mid$ <conditions> $\}$, where <conditions> is some expression that defiens the conditions satisfied by all elements x in this subset. The vertical bar symbol "$|$" is read "such that," so the whole set-builder statement reads as "the set of numbers x in S such that <conditions>." For example, the interval $[a, b] \subset \mathbb{R}$ is defined as $[a, b] \stackrel{\text{def}}{=} \{x \in \mathbb{R} \mid a \leqslant x \leqslant b\}$, where the condition $a \leqslant x \leqslant b$ describes precisely the numbers between a and b, inclusively. The interval $[a, b) \subset \mathbb{R}$, is defined as $[a, b) \stackrel{\text{def}}{=} \{x \in \mathbb{R} \mid a \leqslant x < b\}$, which reads "the set of real numbers that are greater than or equal to a and strictly less than b."

Functions We use the notation $f : A \to B$ to describe a function f that takes inputs from the set A and produces outputs in the set B. The most common type of function has the form $f : \mathbb{R} \to \mathbb{R}$, which means it is a function that takes real numbers as inputs and produces real numbers as outputs. We often denote the function input as x and the function output as $f(x)$ or y. The following table shows some examples of commonly used functions.

Definition	Name	Used to denote
$f(x) \stackrel{\text{def}}{=} mx + b$	linear	the line with slope m and initial value b
$f(x) \stackrel{\text{def}}{=} x^2$	quadratic	the square of x
$f(x) \stackrel{\text{def}}{=} \sqrt{x}$	square root	the square root of x
$f(x) \stackrel{\text{def}}{=} \lvert x \rvert$	absolute value	the absolute value of x
$f(x) \stackrel{\text{def}}{=} e^x$	expnential	the exponential function base e
$f(x) \stackrel{\text{def}}{=} \ln(x)$	natural log of x	the logarithm base e
$f(x) \stackrel{\text{def}}{=} 10^x$	power of 10	the exponential function base 10
$f(x) \stackrel{\text{def}}{=} \log_{10}(x)$	log base 10 of x	the logarithm base 10

You can think of these functions as basic building blocks that can be transformed and combined to describe arbitrary relationships between an input variable x and the output variable $y = f(x)$.

We use the notation $f(x)$, read "f of x," to describe the output of the function f applied to the input x. For example, if we define the function f as $f(x) \stackrel{\text{def}}{=} 3x + 5$, anytime we see the expression $f(a)$, we know this expression is equivalent to the function's output: $3a + 5$.

Whenever we're solving an equation involving the function $f(x)$, it is very useful to know the *inverse function* of the function f, which we denote f^{-1}. The inverse function f^{-1} acts as the "undo" operation for the function f. If you apply f^{-1} to the output of f, you get back the original x you started from: $f^{-1}(f(x)) = x$. For example, the inverse of the function $f(x) \stackrel{\text{def}}{=} 3x + 5$ is the function $f^{-1}(x) =$

$\frac{1}{3}(x-5)$. Try choosing any number $x = a$ and calculate $y = f(x)$, then calculate $f^{-1}(y)$ to verify that you get back to the value a you started from. Here are some other examples of functions and their inverses. The inverse of the square root function $g(x) \stackrel{\text{def}}{=} \sqrt{x}$, is the quadratic function: $g^{-1}(x) = x^2$. The inverse of the exponential $h(x) = e^x$, is logarithm base e: $h^{-1}(x) = \ln(x)$. An inverse relationship also exists between the exponential function base 10, and the logarithm base 10.

* * *

If this is the first time you're seeing the math concepts and notation we presented the previous paragraphs, you might benefit from a more through review of high school math concepts. I'd like to use this opportunity to shamelessly plug the *No Bullshit Guide to Mathematics*[Sav18], since it is precisely on topic.

Data notation

We denote a sample of size n using a boldface symbol $\mathbf{x} = [x_1, x_2, x_3, \ldots, x_n]$. You can think of $\mathbf{x}$ as the measurements of the variable x collected from n individuals. The sample $\mathbf{x}$ will often be stored as a column named x in a Pandas data frame.

For a sample of numerical values, we can compute the following *descriptive statistics* to capture its essential characteristics.

Symbol	Read as	Denotes
n	sample size	number of observations in the sample $\mathbf{x}$
Mean	mean	average value
Med	median	middle value of the dataset
Mode	mode	the most frequently observed value
Var	variance	average squared deviation from the mean
Std	standard deviation	the square root of the variance
$\mathbf{Q}_1$	first quartile	one quarter of values smaller than $\mathbf{Q}_1$
$\mathbf{Q}_2$	second quartile	middle value of the sample $\mathbf{x}$
$\mathbf{Q}_3$	third quartile	one quarter of values are larger than $\mathbf{Q}_3$
IQR	interquartile range	span of the middle fifty percent of the data
Min	minimum	the smallest value in the sample $\mathbf{x}$
Max	maximum	the largest value in the sample $\mathbf{x}$
Range	range	difference between the **Max** and **Min**

You should think of the descriptive statistics as functions you compute from the sample $\mathbf{x}$. For example, the minimum value in the sample $\mathbf{x}$ is denoted $\mathbf{Min}(\mathbf{x})$. Refer to Table **??** (page **??**) to see Pandas methods we use compute descriptive statistics. The sample mean, variance, and standard deviation are the most common statistics, so we use the special shorthand notation for them: $\bar{\mathbf{x}} = \mathbf{Mean}(\mathbf{x})$, $s_{\mathbf{x}}^2 = \mathbf{Var}(\mathbf{x})$, and $s_{\mathbf{x}} = \mathbf{Std}(\mathbf{x})$.

Bivariate dataset A bivariate dataset consists of pairs of observations from two variables, $[\mathbf{x}, \mathbf{y}] = [(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)]$. The *covariance* is denoted $\mathbf{Cov}(\mathbf{x}, \mathbf{y})$ and it measures the joint variability of $\mathbf{x}$ and $\mathbf{y}$. Another descriptive statistics is the *correlation*, which is denoted $\mathbf{Corr}(\mathbf{x}, \mathbf{y})$ and measures the degree of linear relatedness between $\mathbf{x}$ and $\mathbf{y}$.

Categorical data Consider now a sample $\mathbf{x} = [x_1, x_2, x_3, \ldots, x_n]$ of categorical values. The frequency of the value v in the data $\mathbf{x}$ is denoted $\mathbf{Freq}_v(\mathbf{x})$, and is the number of vs in the data $\mathbf{x}$. The *relative frequency* is denoted $\mathbf{RelFreq}_v(\mathbf{x})$, and counts the number of vs in a the data $\mathbf{x}$ divided by n.

Calculus notation

Expression	Denotes
$f(x)$	a function of the form $f : \mathbb{R} \to \mathbb{R}$
$f'(x)$	derivative of $f(x)$
$\frac{d}{dx}$	derivative operator: $\frac{d}{dx}[f(x)] = f'(x)$
$\int_a^b f(x)\,dx$	integral of $f(x)$ between $x = a$ and $x = b$
$F(x)$	the integral function of $f(x)$
a_k	sequence $a_k : \mathbb{N} \to \mathbb{R}$, also denoted $(a_0, a_1, a_2, a_3, \dots)$
$\displaystyle\sum_{k=r}^{s} a_k$	the summation $a_r + a_{r+1} + a_{r+2} + \cdots + a_{s-1} + a_s$

TODO: condensed summary of calc. with examples

Probability notation

Expression	Denotes
$\{\text{descr}\}$	an outcome described by the condition "descr"
$\Pr(\{\text{descr}\})$	the probability of the outcome $\{\text{descr}\}$
$\Pr(\{\text{descr}\}\mid\{\text{cond}\})$	conditional probability of outcome $\{\text{descr}\}$ given $\{\text{cond}\}$
$\mathbb{1}_{\{\text{descr}\}}$	indicator function for the outcome $\{\text{descr}\}$
X	a random variable (henceforth abbreviated as r.v.)
$\mathcal{X}$	the set of possible outcomes for the r.v. X
x	a particular outcome of the r.v. X
$f_X(x) \stackrel{\text{def}}{=} \Pr(\{X = x\})$	*probability mass function* (pmf) of a discrete r.v. X, or *probability density function* (pdf) of a continuous r.v. X.
$X \sim f_X$	X is distributed according to f_X
$F_X(x) \stackrel{\text{def}}{=} \Pr(\{X \leqslant x\})$	cumulative distribution function (CDF) of the r.v. X
$\mathbb{E}_X[\cdot]$	expectation operator with respect to r.v. X
$\mu \stackrel{\text{def}}{=} \mathbb{E}_X[X]$	the *mean* of X
$\sigma^2 \stackrel{\text{def}}{=} \mathbb{E}_X[(X-\mu)^2]$	the *variance* of X; also denoted $\mathbf{var}(X)$
$\sigma = \sqrt{\sigma^2}$	the *standard deviation* of X
$\mathcal{N}(\mu,\sigma)$	the *normal* distribution with mean μ and standard deviation σ. The probability density function for the r.v. $X \sim \mathcal{N}(\mu,\sigma)$ is $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$
$Z \sim \mathcal{N}(0,1)$	the *standard normal distribution*
$\Gamma(z)$	the Gamma function $\Gamma(z) \stackrel{\text{def}}{=} \int_{y=0}^{y=\infty} y^{z-1}e^{-y}dy$
$\mathbf{X} \stackrel{\text{def}}{=} (X_1, X_2, \ldots, X_n)$	random sample of size n. Each $X_i \sim f_X$
$\mathbb{E}_{\mathbf{X}}[\cdot]$	expectation with respect to random sample $\mathbf{X}$
$\mathbf{x} \stackrel{\text{def}}{=} (x_1, x_2, \ldots, x_n)$	sample of n observations
$f_{\mathbf{x}}$	empirical pmf of sample $\mathbf{x} = (x_1, x_2, \ldots, x_n)$
$F_{\mathbf{x}}$	empirical CDF of sample $\mathbf{x} = (x_1, x_2, \ldots, x_n)$
$F_{\mathbf{x}}^{-1}$	inverse of the empirical CDF of the sample $\mathbf{x}$

TODO: condensed summary of prob. concepts with examples

Probability distributions reference

Name	Math notation	Code	Parameters	Notes
Discrete uniform	$\mathcal{U}_d$	`randint(alpha, beta+1)`	$\alpha, \beta \in \mathbb{Z}$	
Bernoulli	Bernoulli	`bernoulli(p)`	$p \in [0,1]$	
Binomial	Binom	`binom(n,p)`	$n \in \mathbb{N}_+$, $p \in [0,1]$	
Poisson	Pois	`poisson(lam)`	$\lambda \in \mathbb{R}_+$	
Geometric	Geom	`geom(p)`	$p \in (0,1)$	
Negative binomial	NBinom	`nbinom(r,p)`	$r \in \mathbb{N}_+$, $p \in (0,1)$	(optiona
Hypergeometric	Hypergeom	`hypergeom(a+b,a,n)`	$a, b, n \in \mathbb{N}_+$	(optiona
Multinomial	Multinomial	`multinomial`	$n \in \mathbb{N}_+$, $p_i \in [0,1]$	(optiona
Uniform	$\mathcal{U}$	`uniform`	$\alpha, \beta \in \mathbb{R}$	
Exponential	Expon	`expon`	$\lambda \in \mathbb{R}_+$	
Normal	$\mathcal{N}$	`norm`	$\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}_+$	...
Standard normal	$Z \sim \mathcal{N}(0,1)$	`norm(0,1)`		...
Gamma	Gamma	`gamma`		
Beta	Beta	`beta`		
Student's T	$\mathcal{T}$	`t`		
Chi square	χ^2	`chi2`		
Snedecor's F	$\mathcal{F}$	`f`		
Cauchy	Cauchy	`cauchy`		(optiona
Laplace	Laplace	`laplace`		(optiona
Pareto	Pareto	`pareto`		(optiona
Wald	Wald	`wald`		(optiona
Weibull	Weibull	`weibull_min`		(optiona
Zipf	Zipf	`zipf`		(optiona

Each of these models provides you with the same set of methods for computing probabilities, confidence intervals, and generating random values: `.pmf(x)` or `.pdf(x)`, `.cdf(b)`, `.expect(f)`, `.mean()`, `.std()`, `.rvs(n)`, etc.

TODO: show math params for all distributions, and equivalent params for computer models `scipy.stats`

Statistics notation

Expression	Denotes
f_X	*probability mass function* of a discrete r.v. X, or *probability density function* of a continuous r.v. X.
n	sample size
$\mathbf{x} \stackrel{\text{def}}{=} (x_1, x_2, \ldots, x_n)$	a particular sample of size n from the r.v. X
$\mathbf{X} \stackrel{\text{def}}{=} (X_1, X_2, \ldots, X_n)$	random sample of size n. Each $X_i \sim f_X$
$\hat{\theta} = g(\mathbf{x})$	particular value of the estimator g computed from the sample $\mathbf{x} = (x_1, x_2, \ldots, x_n)$
$\widehat{\Theta} = g(\mathbf{X})$	sampling distribution of the estimator g computed from the random sample $\mathbf{X} = (X_1, X_2, \ldots, X_n)$
$\overline{\mathbf{x}}$	sample mean
$s_{\mathbf{x}}^2$	sample variance
$s_{\mathbf{x}}$	sample standard deviation
$\mathbf{se}_g$	standard error of the estimator g
$\widehat{\mathbf{se}}_g$	estimated standard error of the estimator g
θ	a parameter of the probability distribution
$\hat{\theta}$	an estimate of the parameter θ
μ	population mean
σ^2	population variance
σ	population standard deviation
ν	degrees of freedom parameter
CV_α	cutoff value for a hypothesis testing decision

Linear models notation

Expression	Denotes
n	number of observations
$[\mathbf{x}, \mathbf{y}]$	bivariate dataset $[(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)]$
(x_i, y_i)	the i^{th} observation in the dataset
$\mu_Y(x) = \beta_0 + \beta_1 x$	regression equation of the mean
$\mathcal{E} \sim \mathcal{N}(0, \sigma)$	normally-distributed error term
$Y(x) \sim \beta_0 + \beta_1 x + \mathcal{E}$	linear model (simple linear regression)
β_0	intercept parameter
β_1	slope parameter
σ	standard deviation parameter
$\mu_{\widehat{Y}}(x) = \widehat{\beta}_0 + \widehat{\beta}_1 x$	estimated regression equation of the mean
$\widehat{Y}(x) \sim \widehat{\beta}_0 + \widehat{\beta}_1 x + \mathcal{N}(0, \widehat{\sigma})$	estimated linear model
$\widehat{\beta}_0$	estimated intercept parameter (denoted `b0` in code)
$\widehat{\beta}_1$	estimated slope parameter (denoted `b0` in code)
$\widehat{\sigma}$	estimated standard deviation parameter
$\widehat{y}_i = \widehat{\beta}_0 + \widehat{\beta}_1 x_i$	fitted values (model predictions for the dataset)
$r_i = y_i - \widehat{y}_i$	residuals (prediction errors)
x_{new}	new, previously unseen input
$\widehat{y}_{\text{new}} = \widehat{\beta}_0 + \widehat{\beta}_1 x_{\text{new}}$	predicted output value for the input x_{new}
$[\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_p, \mathbf{y}]$	multivariate dataset
$(x_{i1}, x_{i2}, \ldots, x_{ip}, y_i)$	the i^{th} observation in multiple regression model
x_{ik}	value of the k^{th} predictor in the i^{th} observation
β_k	the slope associated with the predictor x_k
SSR	sum of squared residuals $\mathbf{SSR} = \sum_{i=1}^{n}(y_i - \widehat{y}_i)^2$
ESS	explained sum of squares $\mathbf{ESS} = \sum_{i=1}^{n}(\widehat{y}_i - \overline{\mathbf{y}})^2$
TSS	total sum of squares $\mathbf{TSS} = \sum_{i=1}^{n}(y_i - \overline{\mathbf{y}})^2$
R^2	coefficient of determination $R^2 = \frac{\mathbf{ESS}}{\mathbf{TSS}} = 1 - \frac{\mathbf{SSR}}{\mathbf{TSS}}$
$\widehat{\mathbf{se}}_{\mu_{\widehat{Y}}}(x)$	uncertainty in the prediction of the mean μ_Y
$\mathbf{ci}_{\mu_Y,\gamma}(x)$	γ-confidence interval for the mean
$\widehat{\mathbf{se}}_{\widehat{y}}(x)$	uncertainty in the prediction of the value Y
$\mathbf{ci}_{Y,\gamma}(x)$	γ-confidence interval for observations