

# Codebook for the data frame representing information about a collection of emails

## Usage

This data is available in the `openintro` package as the data frame `email`.

## Description

These data represent incoming emails for the first three months of 2012 for David Diez's Gmail Account, early months of 2012. All personally identifiable information has been removed.

## Format

A data frame with 3921 observations on the following 21 variables.

Variable	Description
<code>spam</code>	Indicator for whether the email was spam.
<code>to_multiple</code>	Indicator for whether the email was addressed to more than one recipient.
<code>from</code>	Whether the message was listed as from anyone (this is usually set by default for regular outgoing email).
<code>cc</code>	Indicator for whether anyone was CCed.
<code>sent_email</code>	Indicator for whether the sender had been sent an email in the last 30 days.
<code>time</code>	Time at which email was sent.
<code>image</code>	The number of images attached.
<code>attach</code>	The number of attached files.
<code>dollar</code>	The number of times a dollar sign or the word “dollar” appeared in the email.
<code>winner</code>	Indicates whether “winner” appeared in the email.
<code>inherit</code>	The number of times “inherit” (or an extension, such as “inheritance”) appeared in the email.
<code>viagra</code>	The number of times “viagra” appeared in the email.
<code>password</code>	The number of times “password” appeared in the email.
<code>num_char</code>	The number of characters in the email, in thousands.
<code>line_breaks</code>	The number of line breaks in the email (does not count text wrapping).
<code>format</code>	Indicates whether the email was written using HTML (e.g. may have included bolding or active links).
<code>re_subj</code>	Whether the subject started with “Re:”, “RE:”, “re:”, or “rE:”
<code>exclaim_subj</code>	Whether there was an exclamation point in the subject.
<code>urgent_subj</code>	Whether the word “urgent” was in the email subject.
<code>exclaim_mess</code>	The number of exclamation points in the email message.
<code>number</code>	Factor variable saying whether there was no number, a small number (under 1 million), or a big number.

*This information was copied from ?email on March 20, 2016*