eBook Writing Workshop – Practical 1: A First Logfile-Style eBook

Introduction

In this practical we will aim to get you familiar with both the TREE (Template Reading & Execution Environment) and DEEP (Documents with Embedded Execution & Provenance) interfaces of Stat-JR.

To recap, Stat-JR has a modular system of templates, each defining a certain function (or suite of functions). Some templates fit models, others plot charts, some produce data summaries, and so on. One of the advantages of this system is that Stat-JR's functionality can be extended simply by adding additional template files.

Both TREE and DEEP are operated through a browser (although they can – and for most users will – be running locally on your machine). The TREE interface is a flexible menu-driven point-and-click environment in which you pair-up templates with datasets. DEEP, on the other hand, is Stat-JR's eBook-reading interface: it still uses templates and datasets, but there is greater scope to provide tailored contextual information; however it is a less flexible environment than TREE in that the choice of templates and datasets is typically more circumscribed.

In the TREE interface we will perform some descriptive statistics and some first model-fitting to your dataset. We will also create an eBook, using the eBook-writing functions in TREE, that we will subsequently export and view via Stat-JR's eBook-reading interface, DEEP. As such this practical will introduce you to TREE's eBook-writing functionality, and navigating around DEEP. The executions our eBook performs will be predetermined, but in later practicals we will explore how to make eBooks more interactive: giving the eBook-reader the opportunity to guide the executions which take place in an eBook.

This practical is written using the Junior School Project (JSP) dataset however the idea is that each participant will use their own dataset to perform a similar exploration and analysis.

Preparation for the Practical

We will of course be using Stat-JR in this and other practicals, and so it is important that you have access to Stat-JR. For this workshop we will supply memory sticks that contain the latest (soon to-be-released) version and so it should be simply a case of plugging the stick into the machines and running Stat-JR from the stick.

The Stat-JR main directory has several subdirectories, including one called **templates** and another called **datasets**; here you will find all the templates and datasets that populate the list available from within Stat-JR's TREE interface. In our final practical we will delve into the **templates** subdirectory to explore and modify some of the template code, but for now we will just use them as they are. In order to use your own data you will need to add your dataset (in Stata **.dta** format) to this directory. Note that if you do not use Stata then one route to constructing a **.dta** file is to load your data into MLwiN or SPSS and then save it in **.dta** format.[1] If you have problems then ask for help at this point.

[1] Alternatively, if your data is saved as a **.txt** file, you can use Stat-JR's **LoadTextFile** template to save it into the temporary memory cache; the dataset will then be available for use in the current session, but you will need to

Getting Started with TREE

Although the workshop is primarily about eBooks, and thus the DEEP interface via which Stat-JR eBooks are read, to *construct* eBooks we will use the TREE environment.

To start up the TREE interface, double-click **tree.cmd** in the base directory of the Stat-JR install (on your memory stick); this will bring up a command window in which a list of commands will appear similar to the following:

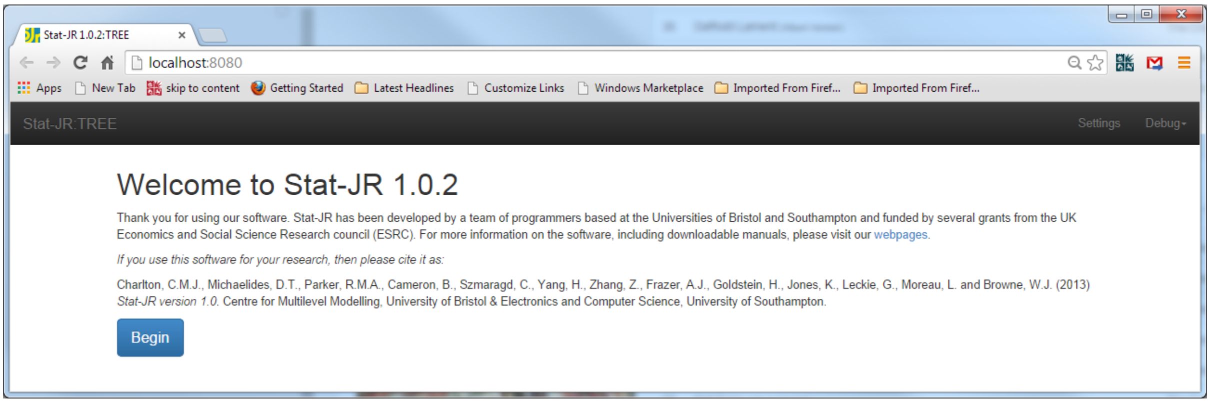
```
E:\newstruct\Software\StatJRrep\estat\trunk>SET Path=E:\newstruct\Software\StatJ
Rrep\estat\trunk\MinGW\bin;C:\Windows\system32;C:\Windows;C:\Windows\System32\Wb
em;C:\Windows\System32\WindowsPowerShell\v1.0\;C:\Program Files (x86)\QuickTime\
QTSystem\;C:\Program Files\TortoiseSVN\bin;C:\Program Files\MiKTeX 2.9\miktex\bi
n\x64\
E:\newstruct\Software\StatJRrep\estat\trunk>SET LTDL_LIBRARY_PATH=E:\newstruct\S
oftware\StatJRrep\estat\trunk\JAGS-3.3.0\i386\modules

E:\newstruct\Software\StatJRrep\estat\trunk>cd src\apps\webtest

E:\newstruct\Software\StatJRrep\estat\trunk\src\apps\webtest>..\..\..\App\Python
.exe webtest.py 8080
WARNING:root:Failed to load package GenStat_model (GenStat not found)
WARNING:root:Failed to load package gretl_model (Gretl not found)
WARNING:root:Failed to load package MATLAB_script (Matlab not found)
WARNING:root:Failed to load package Minitab_model (Minitab not found)
WARNING:root:Failed to load package Minitab_script (Minitab not found)
WARNING:root:Failed to load package MIXREGLS (MIXREGLS not found)
WARNING:root:Failed to load package Octave_script (Octave not found)
WARNING:root:Failed to load package SABRE (Sabre not found)
WARNING:root:Failed to load package SAS_model (SAS not found)
WARNING:root:Failed to load package SAS_script (SAS not found)
WARNING:root:Failed to load package SPSS_model (SPSS not found)
WARNING:root:Failed to load package SPSS_script (SPSS not found)
WARNING:root:Failed to load package Stata_MLwiN (Stata not found)
WARNING:root:Failed to load package Stata_model (Stata not found)
WARNING:root:Failed to load package Stata_script (Stata not found)
WARNING:root:Failed to load package SuperMix (SuperMIX not found)
INFO:root:Trying to locate and open default web browser
http://0.0.0.0:8080/
```

The last line indicates that a web process is starting; Stat-JR uses a web browser as an input/output device however the computation will be done on your own computer. If you haven't got a web browser already open, the default will open and look as follows:

download it (as a **.dta** file) via **Dataset > Download** (e.g. saving it into the **datasets** subdirectory) for use in the future sessions too.



Welcome to Stat-JR 1.0.2

Thank you for using our software. Stat-JR has been developed by a team of programmers based at the Universities of Bristol and Southampton and funded by several grants from the UK Economics and Social Science Research council (ESRC). For more information on the software, including downloadable manuals, please visit our webpages.

If you use this software for your research, then please cite it as:

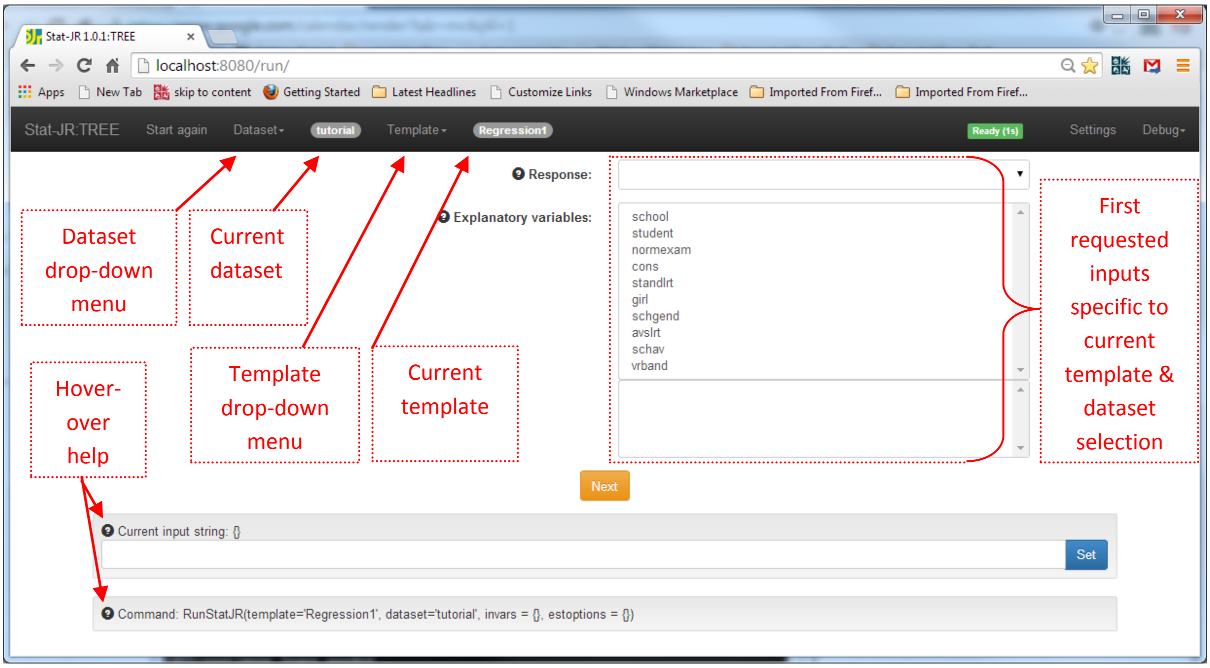
Charlton, C.M.J., Michaelides, D.T., Parker, R.M.A., Cameron, B., Szmaragd, C., Yang, H., Zhang, Z., Frazer, A.J., Goldstein, H., Jones, K., Leckie, G., Moreau, L. and Browne, W.J. (2013) *Stat-JR version 1.0.* Centre for Multilevel Modelling, University of Bristol & Electronics and Computer Science, University of Southampton.

Two important things to note:

- The number 8080 (in this example) will vary each time you run the software to allow several versions of Stat-JR to run at once.

- Stat-JR works best with either Chrome or Firefox, so if the default browser on your machine is Internet Explorer it is best to open a different browser and copy the html path to it.

Clicking on the **Begin** button will then bring up the main screen for Stat-JR, as follows:



Current input string: {}

Command: RunStatJR(template='Regression1', dataset='tutorial', invars = {}, estoptions = {})

- This window shows the current template and current dataset at the top of the screen (in the grey boxes). To the left of each of these is a drop-down menu from which one can select different templates and datasets.

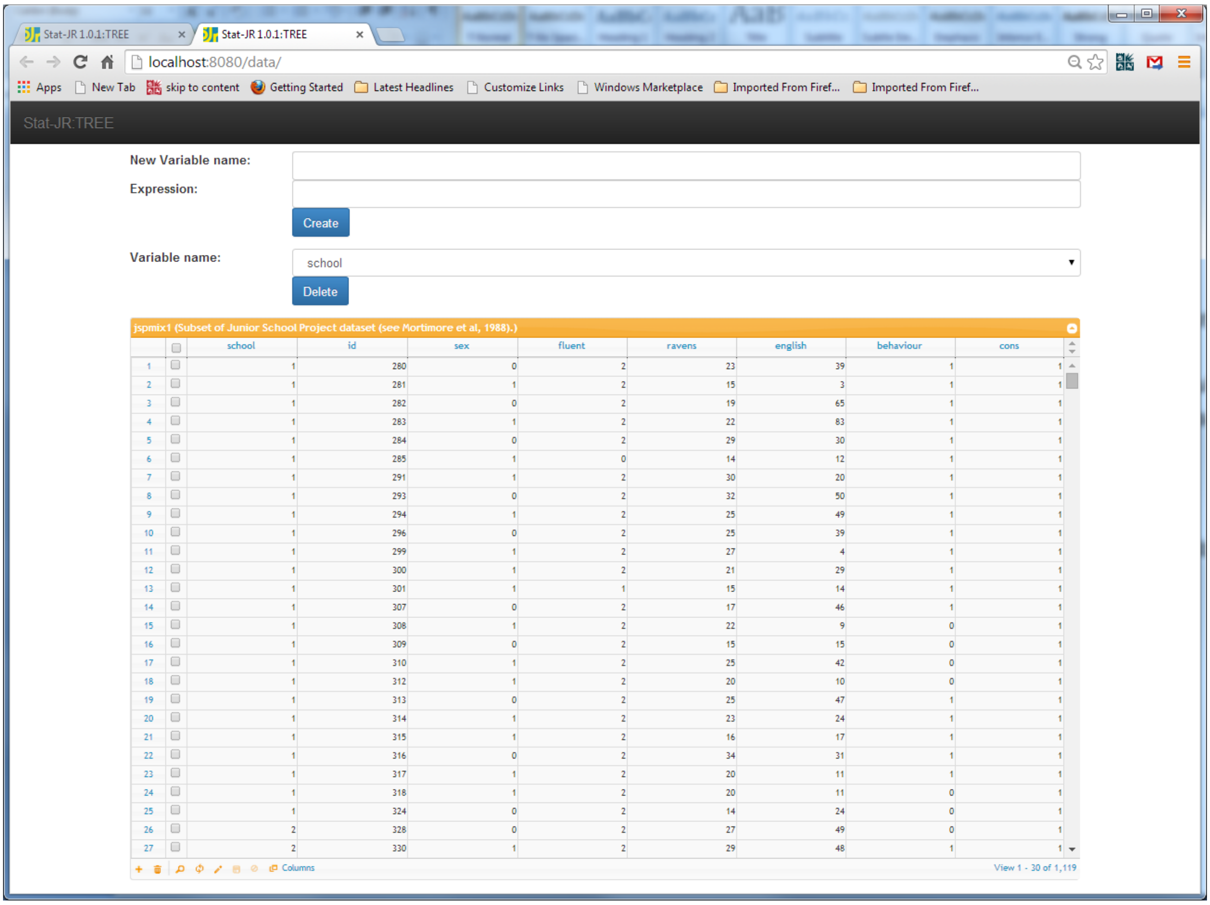
- Underneath the black bar you will see the first input choices specific to the currently-selected template and dataset; these will change as you select different templates and datasets. Here and elsewhere you may see black-circled question marks – hovering your cursor over these will reveal help, as will lingering your cursor over the options in the drop-down lists accessible via the black bar at the top.

- You may notice that the **Current input string** and **Command** boxes, towards the bottom, become further populated as you select your inputs. Inside the curly brackets to the right of the **Current input string** heading will appear a string recording your responses to the inputs above; if you later want to repeat these, without pointing-and-clicking through the input boxes again, you can copy and paste the input string (including the curly brackets) into the white box just below it, and then press **Set** to do so. The **Command** box, on the other hand, writes out a command reflecting your choice of inputs that can then be used in a command-line version of Stat-JR.

Selecting your own dataset

We will first select our own dataset to use in the practical. Here I will use the dataset **jspmix1.dta** but you should use your own dataset. In the black menu bar at the top of the screen you will see a **Dataset** drop-down list to the left of the currently-selected dataset (**tutorial**). Click on the **Dataset** drop-down list and select **Choose.** From the resulting list of datasets scroll down until you find your dataset, in my case **jspmix1**, highlight it, and click on the **Use** button to the bottom right of the window.

After doing so**,** the **Current dataset** will change at the top of the window to confirm your selection, and we can select **View** from the **Dataset** drop-down list which will bring up a separate tab at the top of the screen with the first data records in the dataset displayed, as shown below:



New Variable name:

Expression:

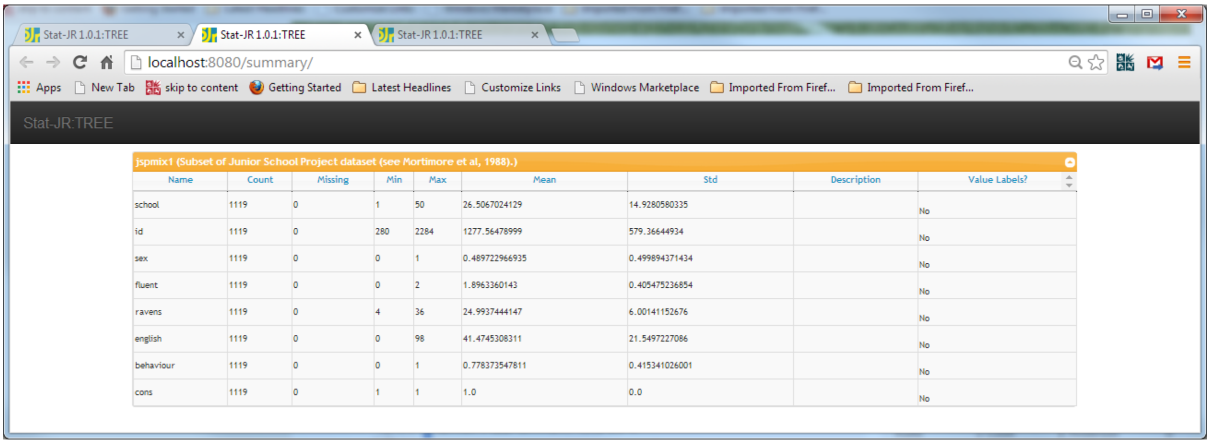
Variable name: school

jspmix1 (Subset of Junior School Project dataset (see Mortimore et al, 1988).)

	□	school	id	sex	fluent	ravens	english	behaviour	cons
1	□	1	280	0	2	23	39	1	1
2	□	1	281	1	2	15	3	1	1
3	□	1	282	0	2	19	65	1	1
4	□	1	283	1	2	22	83	1	1
5	□	1	284	0	2	29	30	1	1
6	□	1	285	1	0	14	12	1	1
7	□	1	291	1	2	30	20	1	1
8	□	1	293	0	2	32	50	1	1
9	□	1	294	1	2	25	49	1	1
10	□	1	296	0	2	25	39	1	1
11	□	1	299	1	2	27	4	1	1
12	□	1	300	1	2	21	29	1	1
13	□	1	301	1	1	15	14	1	1
14	□	1	307	0	2	17	46	1	1
15	□	1	308	1	2	22	9	0	1
16	□	1	309	0	2	15	15	0	1
17	□	1	310	1	2	25	42	0	1
18	□	1	312	1	2	20	10	0	1
19	□	1	313	0	2	25	47	1	1
20	□	1	314	1	2	23	24	1	1
21	□	1	315	1	2	16	17	1	1
22	□	1	316	0	2	34	31	1	1
23	□	1	317	1	2	20	11	1	1
24	□	1	318	1	2	20	11	0	1
25	□	1	324	0	2	14	24	0	1
26	□	2	328	0	2	27	49	0	1
27	□	2	330	1	2	29	48	1	1

View 1 - 30 of 1,119

Here we see the values of the 8 variables in this dataset for the first 27 records. We can get some basic summary information for each column by returning to the other tab and selecting **Summary** from the **Dataset** drop-down list**;** this produces the following in a new tab:



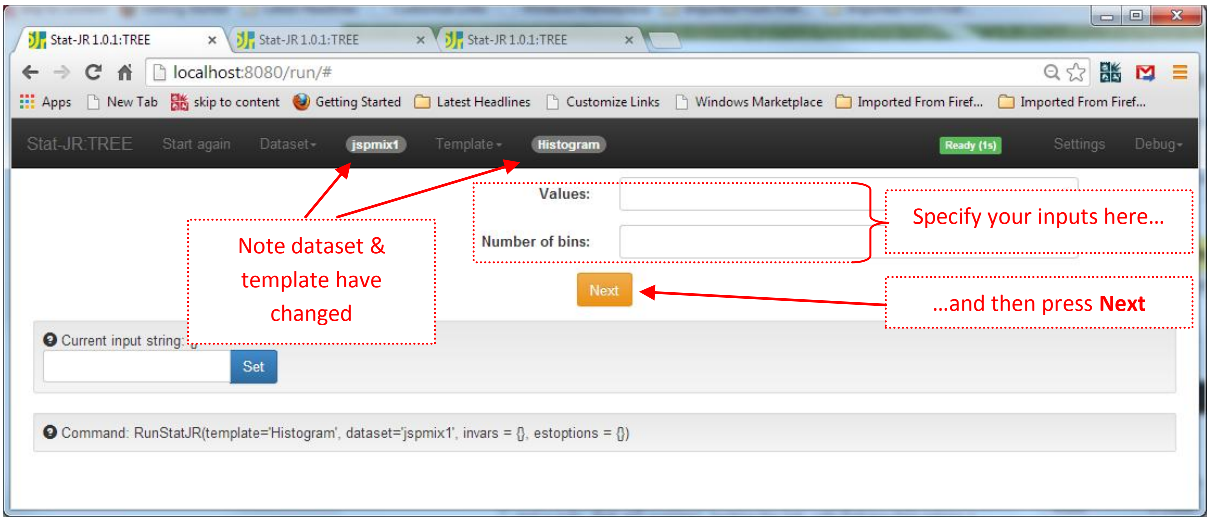
jspmix1 (Subset of Junior School Project dataset (see Mortimore et al, 1988).)

Name	Count	Missing	Min	Max	Mean	Std	Description	Value Labels?
school	1119	0	1	50	26.5067024129	14.9280580335		No
id	1119	0	280	2284	1277.56478999	579.36644934		No
sex	1119	0	0	1	0.489722966935	0.499894371434		No
fluent	1119	0	0	2	1.8963360143	0.405475236854		No
ravens	1119	0	4	36	24.9937444147	6.00141152676		No
english	1119	0	0	98	41.4745308311	21.5497227086		No
behaviour	1119	0	0	1	0.778373547811	0.415341026001		No
cons	1119	0	1	1	1.0	0.0		No

This provides some basic summary statistics for the dataset. In this practical we will look at what factors influence the English scores of the children which are stored in the column **english**.

Looking at the shape of the response variable distribution

We will now choose our first template in Stat-JR, to generate a histogram of our chosen response variable. For the JSP example I have used **english** but you will choose whatever is appropriate in your own dataset. To change the template in Stat-JR return to the main browser tab and from the **Template** drop-down list in the black bar at the top select **Choose**. From the list that appears scroll down until you find **Histogram** (alternatively, you can use the cloud terms to narrow the choice down: e.g. clicking on **Plots**). After selecting **Histogram,** and clicking on the **Use** button, the main tab should look as follows:



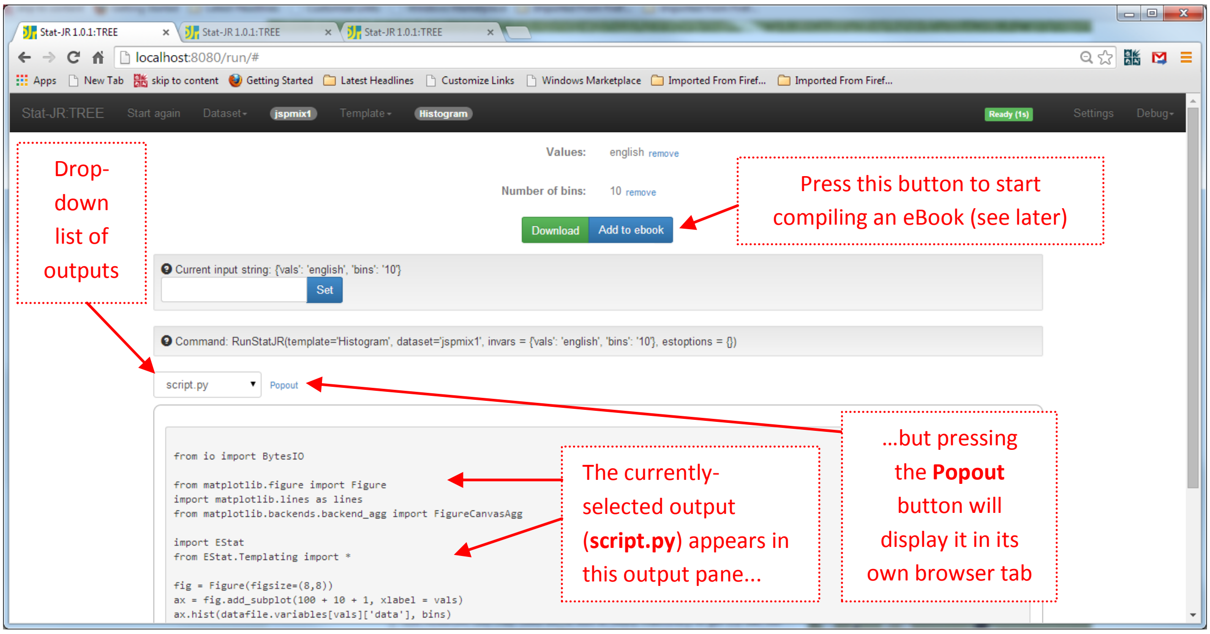
Note dataset & template have changed

Specify your inputs here...

...and then press **Next**

Command: RunStatJR(template='Histogram', dataset='jspmix1', invars = {}, estoptions = {})

Next we need to select the variable to plot (in my case **English** for the box labelled Values) and the number of bins (I'll choose **10**). Then click on **Next** and **Run** and the template will execute and the browser window will look as follows:



Values: english remove

Number of bins: 10 remove

Current input string: {'vals': 'english', 'bins': '10'}

Command: RunStatJR(template='Histogram', dataset='jspmix1', invars = {'vals': 'english', 'bins': '10'}, estoptions = {})

```
from io import BytesIO

from matplotlib.figure import Figure
import matplotlib.lines as lines
from matplotlib.backends.backend_agg import FigureCanvasAgg

import EStat
from EStat.Templating import *

fig = Figure(figsize=(8,8))
ax = fig.add_subplot(100 + 10 + 1, xlabel = vals)
ax.hist(datafile.variables[vals]['data'], bins)
```

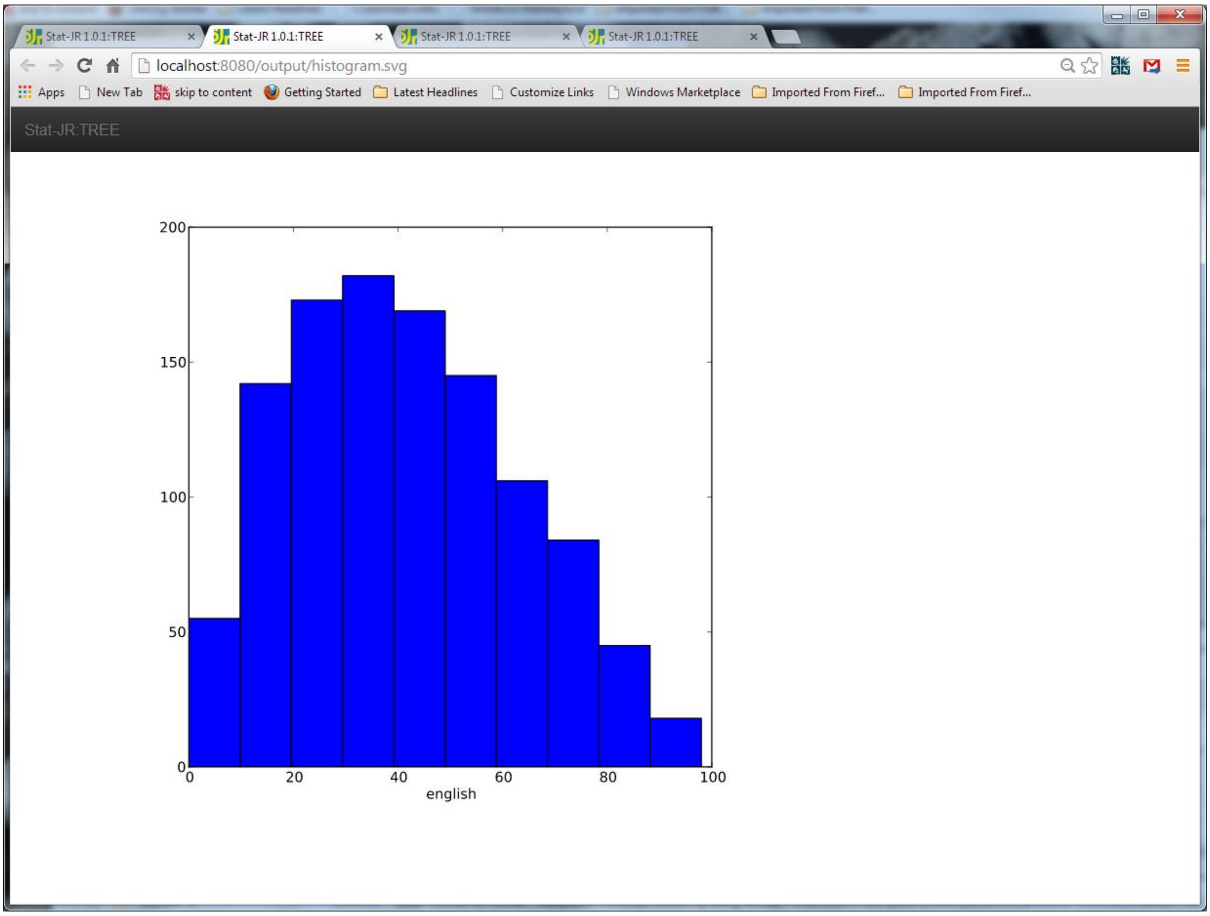
Drop-down list of outputs

Press this button to start compiling an eBook (see later)

The currently-selected output (**script.py**) appears in this output pane...

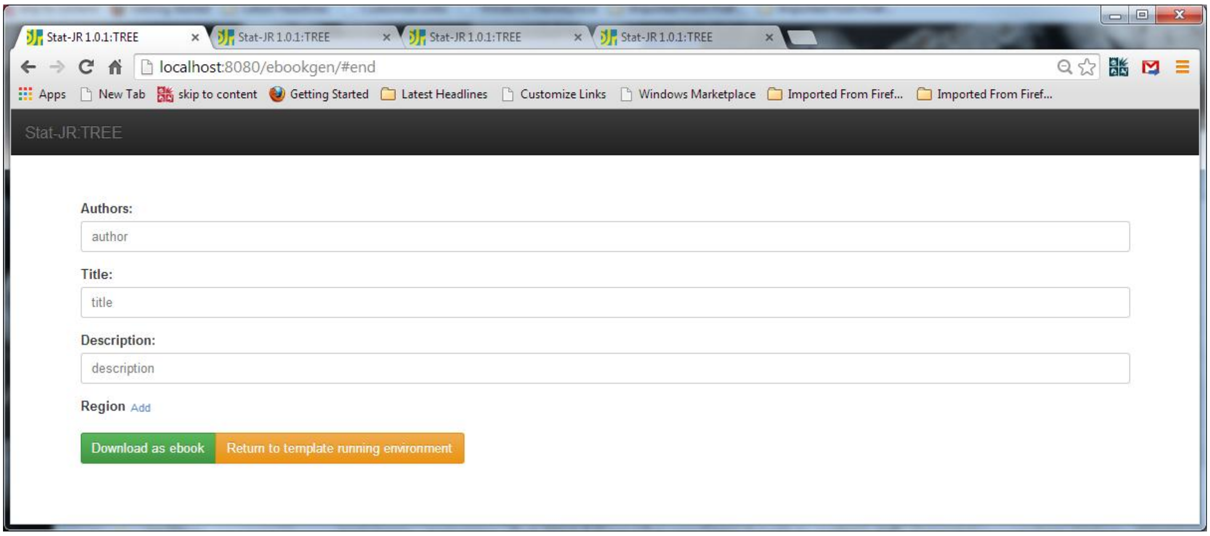
...but pressing the **Popout** button will display it in its own browser tab

Towards the bottom of the screen is a drop-down list of outputs which currently is showing an object called **script.py** which in fact contains the Python code used to create the histogram. We can look at the histogram itself by selecting **histogram.svg** from the pull-down list. Clicking on **Popout** next to the output list will then display the object in its own tab as shown below:



Creating the eBook

Now that we have run our first template we are in a position to start our first eBook. If you return to the main tab you will notice that below the inputs you typed earlier are two buttons. The blue one says **Add to ebook** and this is our route to the eBook writer. Click on the button and the main tab is replaced by the eBook-writer screen thus:

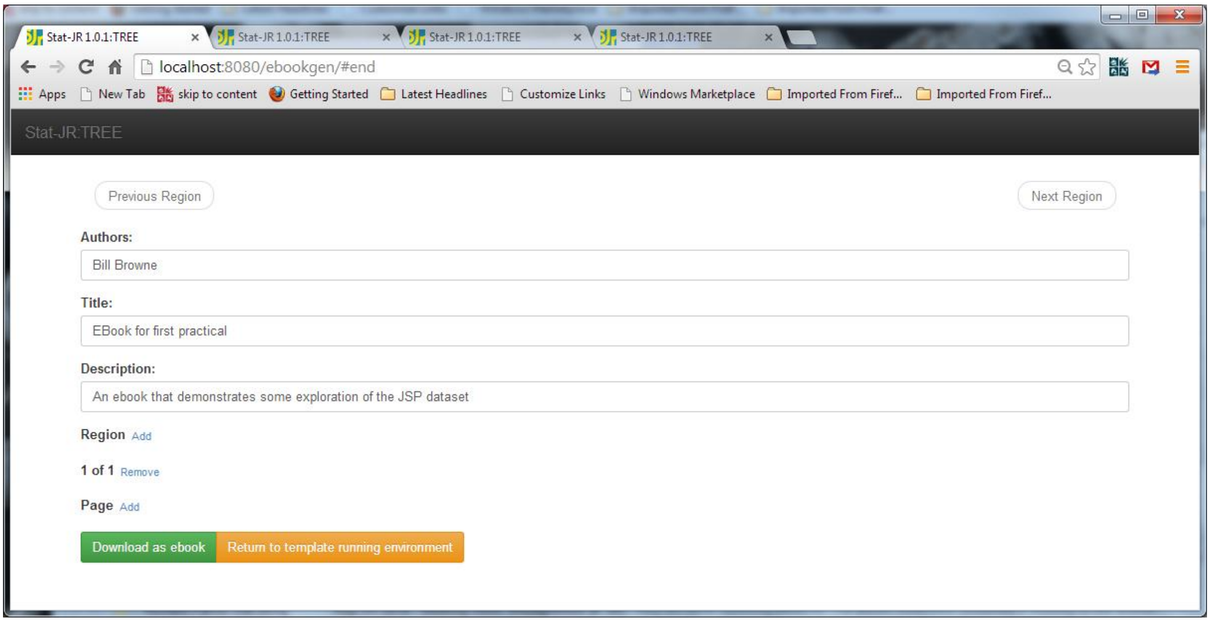


Authors:
author

Title:
title

Description:
description

We can start our eBook by entering some basic information in the top three boxes and then click **Add** next to **Region** to add our first activity region to the eBook. The screen should then look something like this:



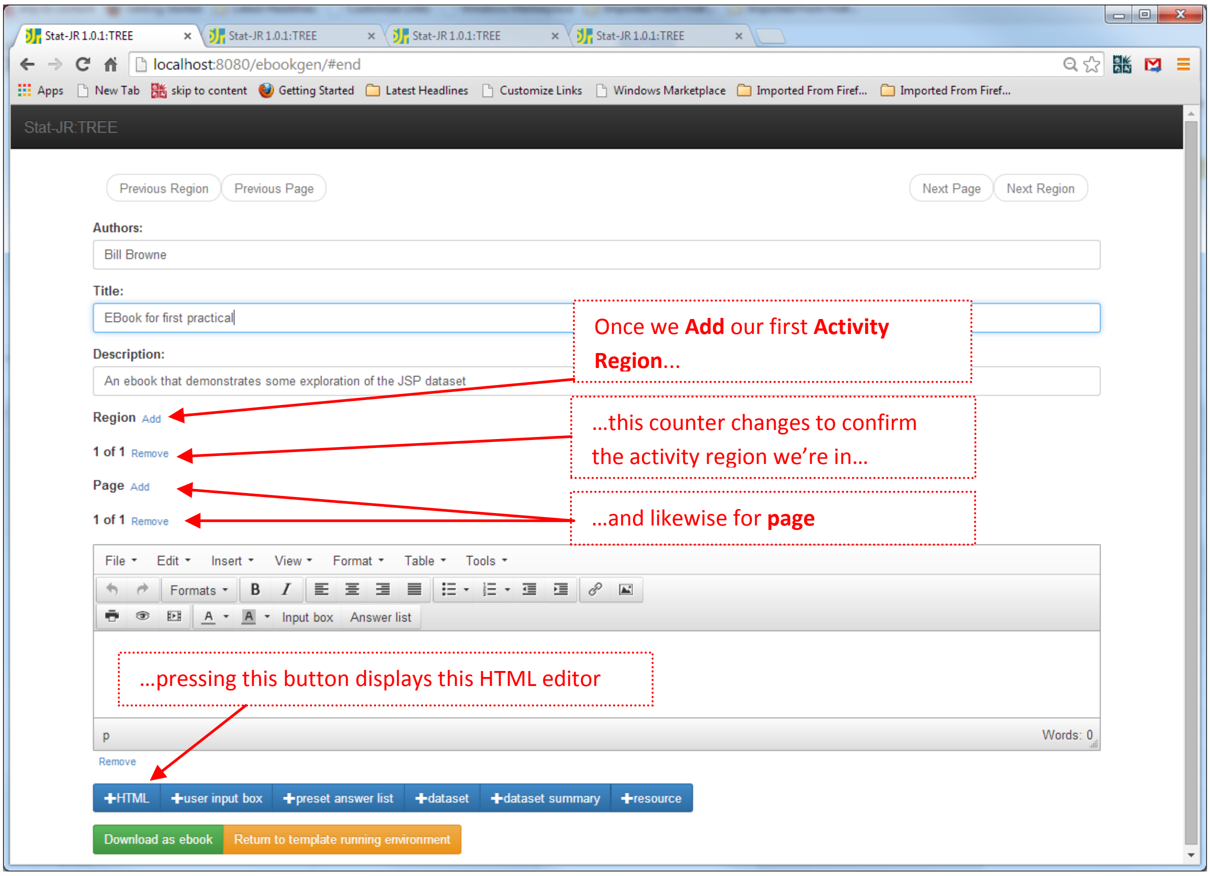
Authors: Bill Browne

Title: EBook for first practical

Description: An ebook that demonstrates some exploration of the JSP dataset

To put this in a little context, an eBook consists of one or more **activity regions**. Each activity region has a circumscribed set of templates, datasets and inputs associated with it. When the eBook user is reading a particular page of an eBook, they will be accessing only one activity region (activity regions can't overlap, and each page of the eBook must only be in one activity region although each activity region can contain several pages). This means that only the objects associated with that particular activity region are referenced, leading to lower computational overheads. Unlike a book, pages in an eBook can be of varying length and so it is largely up to the writer when they decide to start a new page.

Earlier in this practical we looked at a summary of the data and a histogram of one variable so we will attempt to replicate this in our first eBook. We will start by adding our first page to our activity region by clicking on the **Add** button next to **Page** which now gives us (in blue) a palette of objects to add to our eBook. We will first click on **+HTML** to add a HTML box at the start of the page. The HTML boxes are how we add headings and paragraphs of text to our eBook. Our screen now looks as follows:



Authors: Bill Browne

Title: EBook for first practical

Description: An ebook that demonstrates some exploration of the JSP dataset

Region Add

1 of 1 Remove

Page Add

1 of 1 Remove

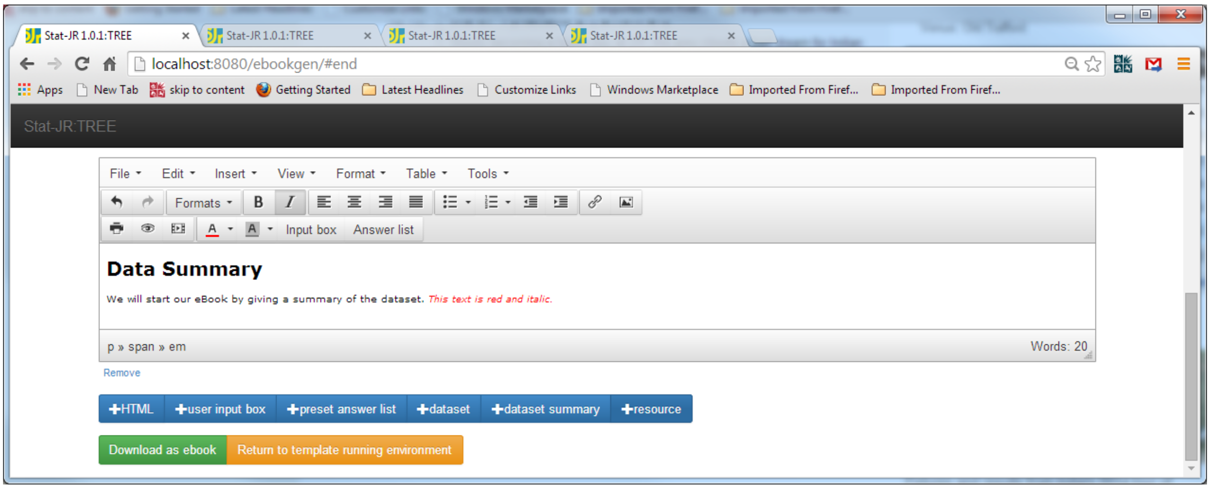
Once we **Add** our first **Activity Region**…

…this counter changes to confirm the activity region we're in…

…and likewise for **page**

…pressing this button displays this HTML editor

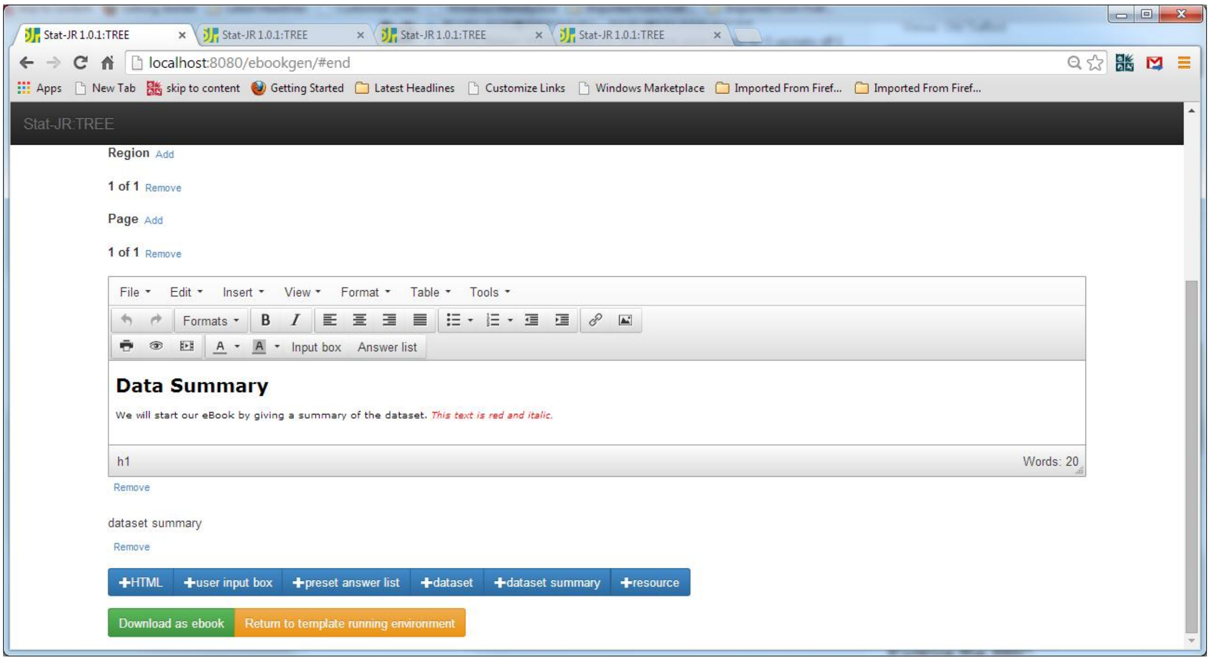
The HTML editor appears above the blue palette and we can now investigate the various options. You will find that there are standard word-processing style options for doing things like making text bold or italicised, changing justification and colour. One of the more important options here is the **Format** menu as this allows the eBook author to specify headings, etc., which are recognised by the eBook reading interface and used to generate hierarchies in its navigation tree. Here I will write a heading and some basic text but feel free to explore yourself:



Data Summary

We will start our eBook by giving a summary of the dataset. *This text is red and italic.*

We now need to add the dataset summary object so click on the **+dataset summary** blue button and the screen will look as follows:



Region Add

1 of 1 Remove

Page Add

1 of 1 Remove

Data Summary

We will start our eBook by giving a summary of the dataset. *This text is red and italic.*

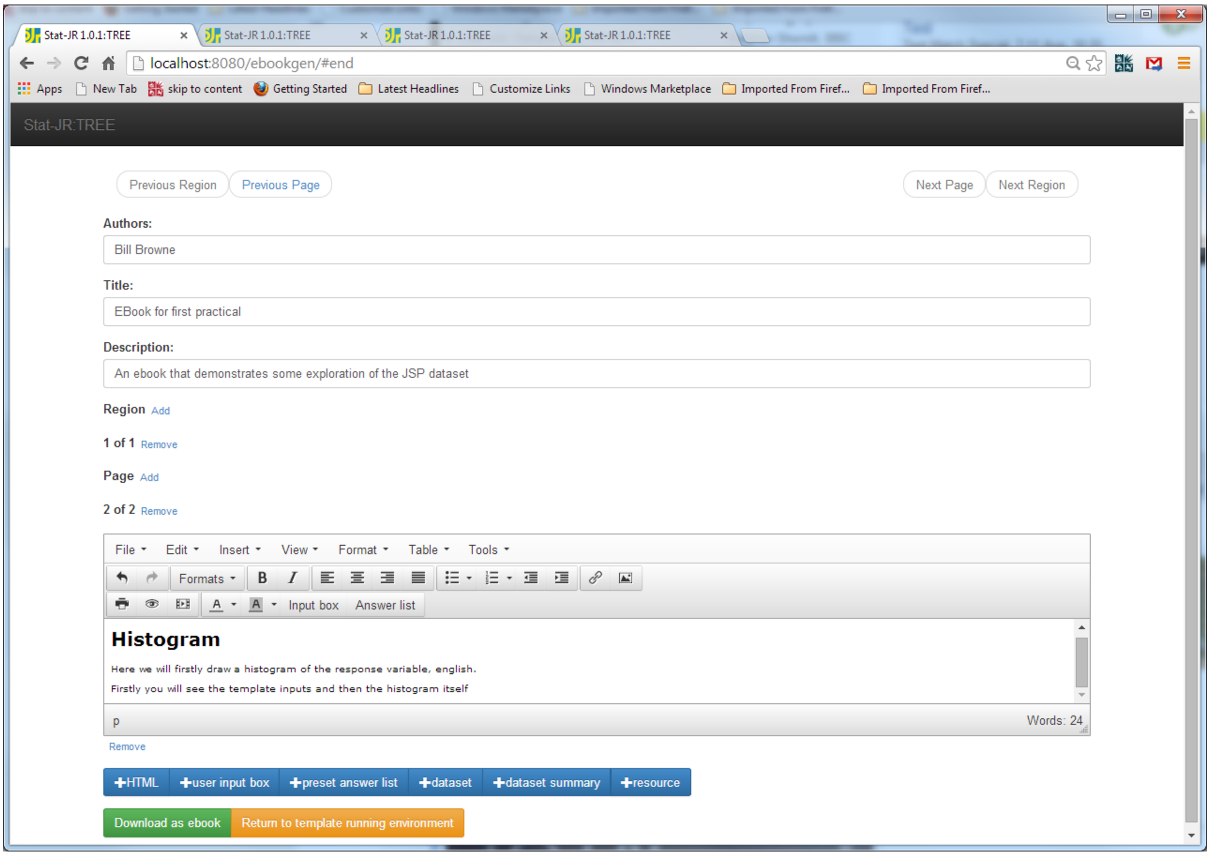
h1 Words: 20

Remove

dataset summary

Remove

Next we will add a second page by clicking on the **Add** button next to **Page**; this will automatically move the focus to the new page and we can add another HTML input here. Again the content of the HTML is somewhat up to you but here is my attempt:



Authors: Bill Browne

Title: EBook for first practical

Description: An ebook that demonstrates some exploration of the JSP dataset

Region Add

1 of 1 Remove

Page Add

2 of 2 Remove

Histogram

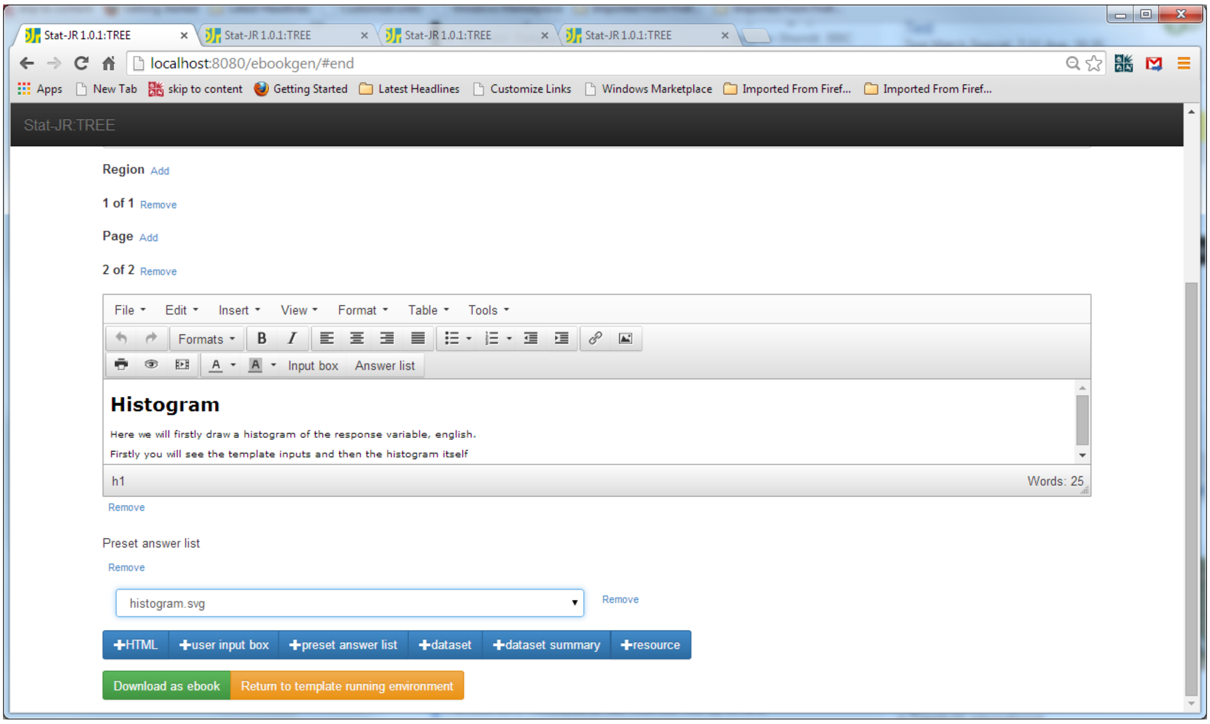
Here we will firstly draw a histogram of the response variable, english.

Firstly you will see the template inputs and then the histogram itself

p Words: 24

Remove

To complete this page we need to add the inputs and the histogram. To do this first click on **+preset answer list** to get a list of inputs appearing and next click on **+resource**. Clicking on **+resource** will invoke a drop-down list from which we can select **histogram.svg**. Depending on your browser and screen resolution, when you hover over the object name in the pull-down list, a preview of the object may appear to the right, which can be useful to help you check your selection. When finished the screen should look as follows:


Stat-JR 1.0.1:TREE
localhost:8080/ebookgen/#end
Stat-JR:TREE
Region Add
1 of 1 Remove
Page Add
2 of 2 Remove
File Edit Insert View Format Table Tools
Formats B I
Input box Answer list
h1
Words: 25
Remove
Preset answer list
Remove
histogram.svg
Remove
+HTML +user input box +preset answer list +dataset +dataset summary +resource
Download as ebook
Return to template running environment

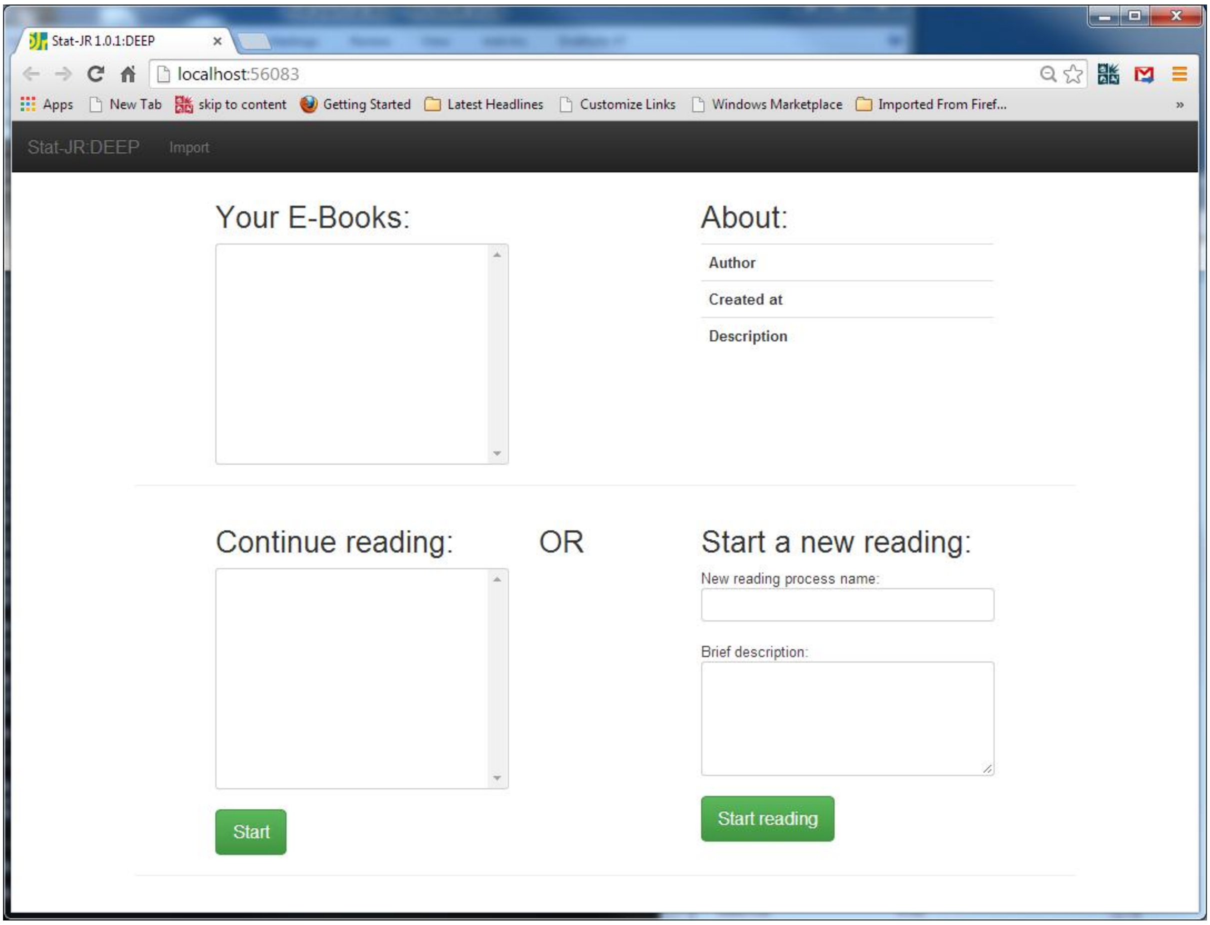
Histogram

Here we will firstly draw a histogram of the response variable, english.
Firstly you will see the template inputs and then the histogram itself

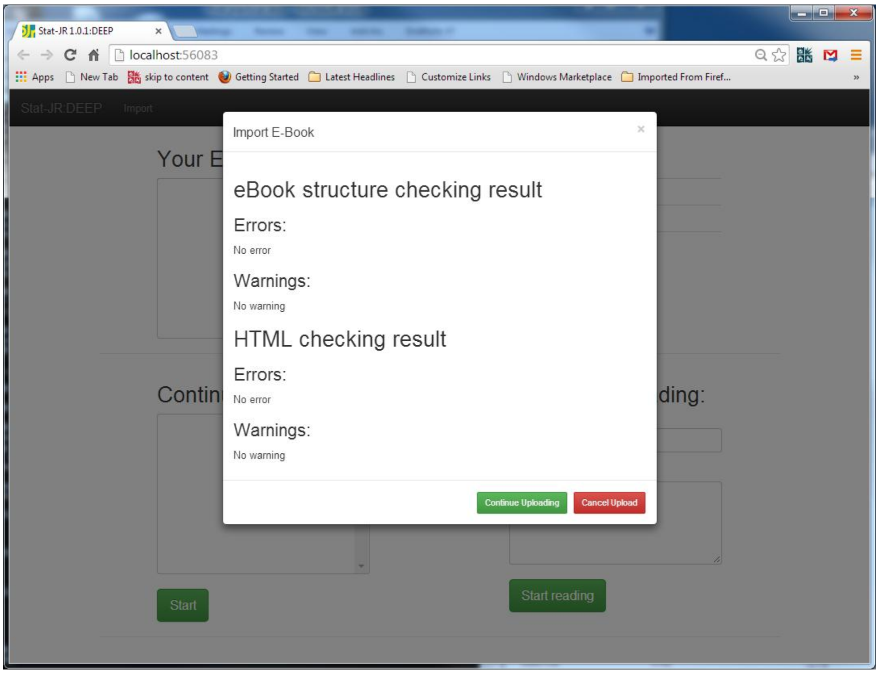
We will eventually add a model to our eBook, but before we do that we'll create a simpler eBook containing the dataset summary and histogram we have just produced. To do this click on the **Download as ebook** button and select a name for the zip file that will contain the eBook. Here I will choose **jsp1.zip** and click **Save** (noting where you have saved the eBook to!) We will next look at this eBook in the DEEP interface, but don't close the TREE browser window as we will return to it soon.

Using DEEP to read our first eBook

The executable for the Stat-JR DEEP interface can be found as **deep.cmd** in the base directory of Stat-JR. Double-clicking on it will bring up another command window and then in the browser will appear the main DEEP window as follows:



As you'll see, we haven't yet loaded any eBooks but if you have used DEEP before the **Your E-Books** list will contain previously loaded eBooks. We want to find our new eBook which I have called *jsp1.zip*. To do this click on **Import** in the black bar at the top, and in the dialogue box that appears click on **+Select an E-Book file** and find the zip file. The system will then check the eBook as shown:


Import E-Book

eBook structure checking result

Errors:
No error

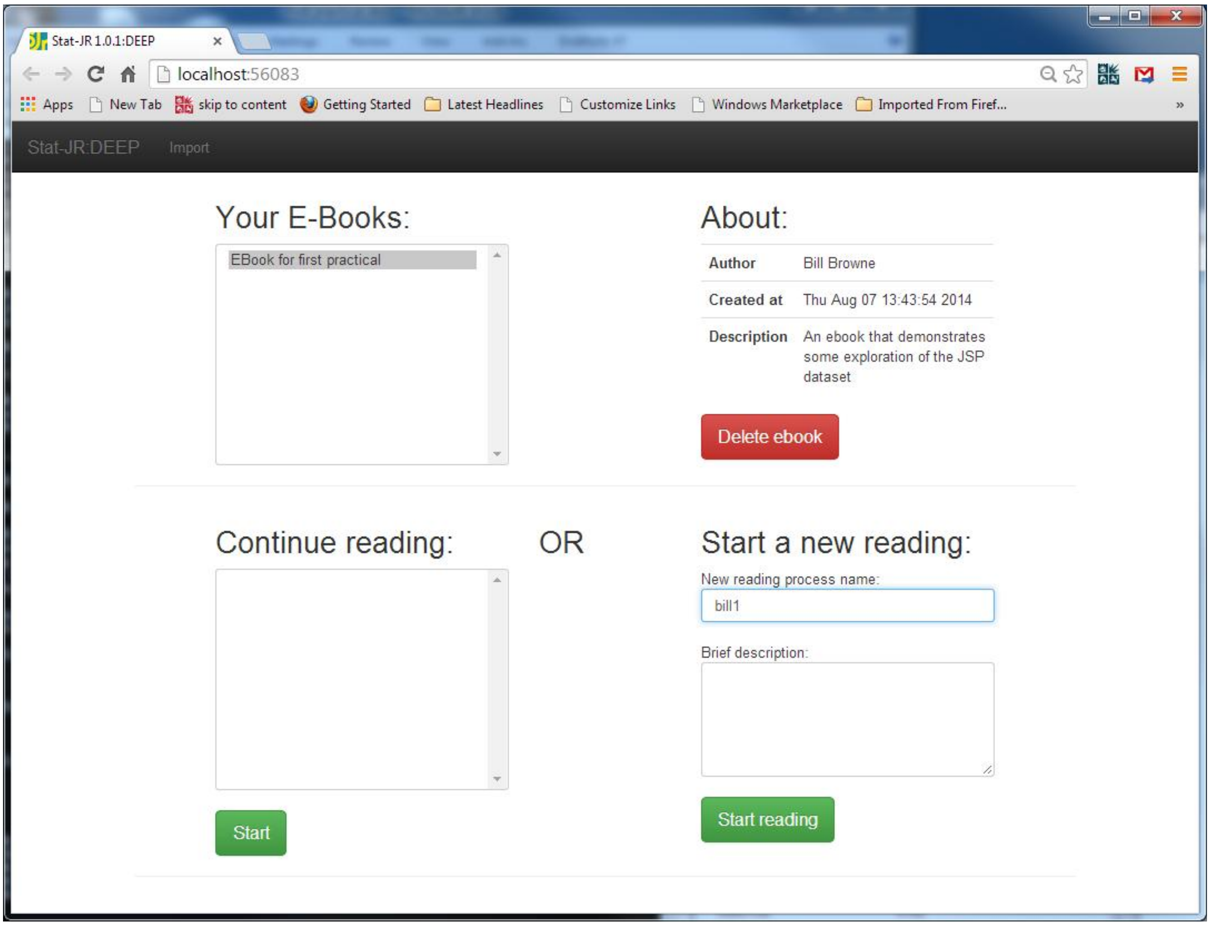
Warnings:
No warning

HTML checking result

Errors:
No error

Warnings:
No warning

If you click on **Continue Uploading** then you will get back to the main screen and the list of eBooks will include this new eBook. Clicking on it, we then need to add a reading process name to uniquely identify this reading of the book (here I've chosen **bill1**) to identify this specific reading of the eBook (and optionally a description) thus:



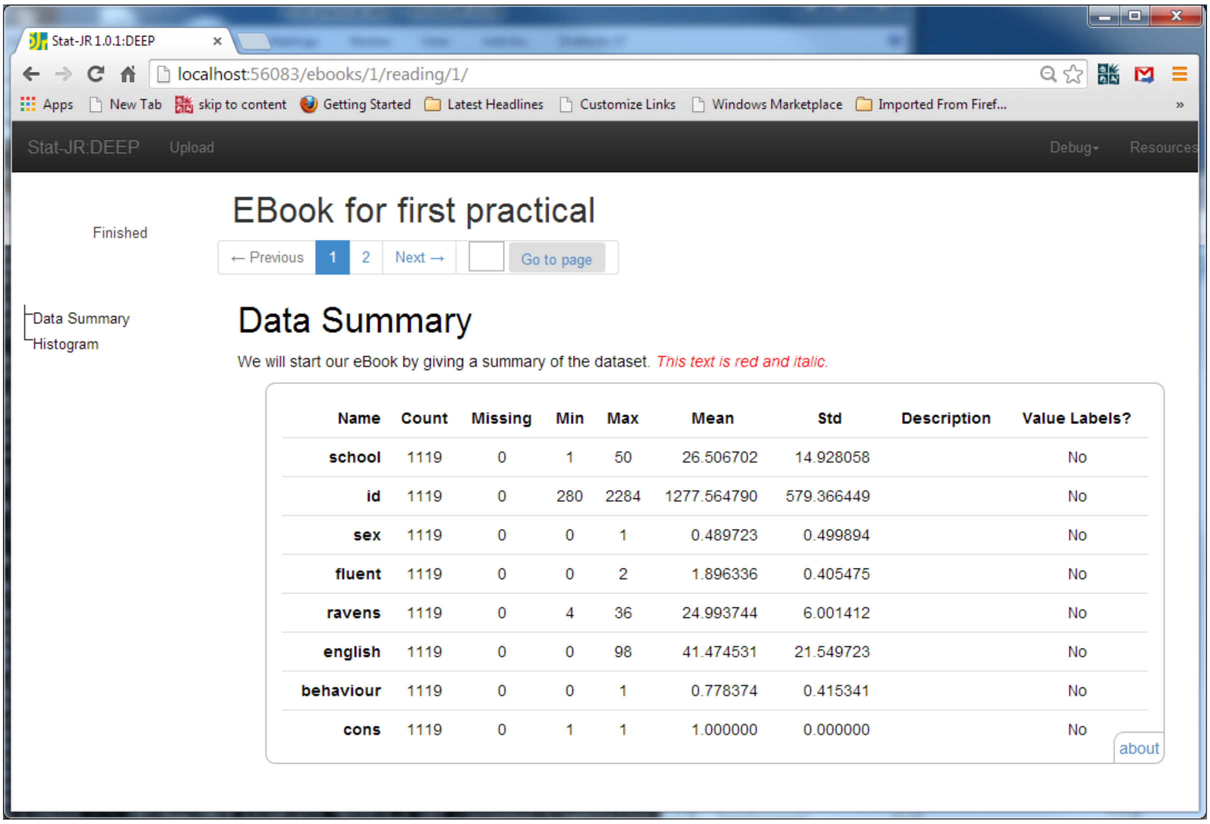
EBook for first practical

Author	Bill Browne
Created at	Thu Aug 07 13:43:54 2014
Description	An ebook that demonstrates some exploration of the JSP dataset

New reading process name: bill1

Brief description:

If we click on **Start reading** the system will go to the first page of the first activity region and start executing the templates within this activity region. Below we can see page 1 of the eBook:



EBook for first practical

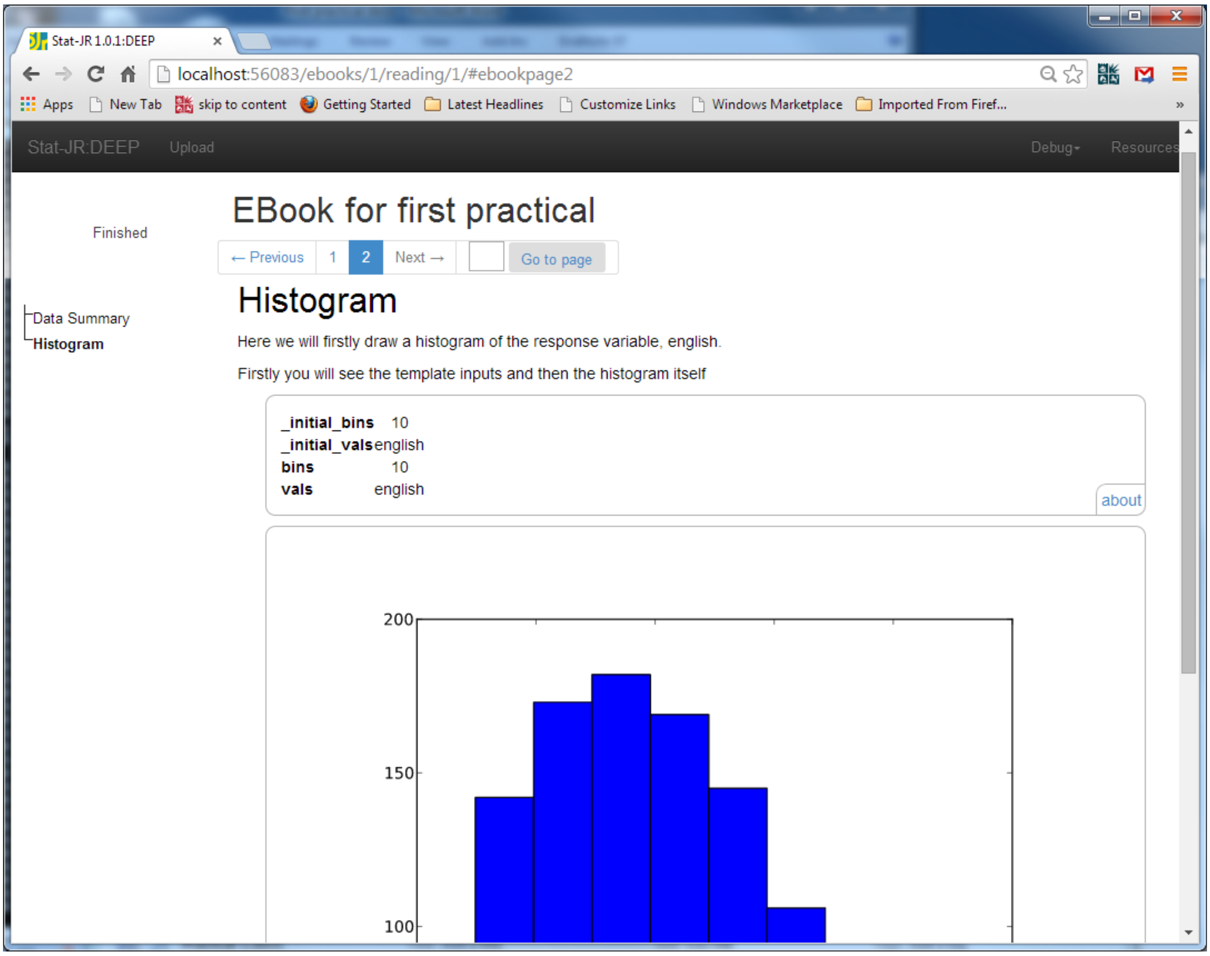
Data Summary

We will start our eBook by giving a summary of the dataset. *This text is red and italic.*

Name	Count	Missing	Min	Max	Mean	Std	Description	Value Labels?
school	1119	0	1	50	26.506702	14.928058		No
id	1119	0	280	2284	1277.564790	579.366449		No
sex	1119	0	0	1	0.489723	0.499894		No
fluent	1119	0	0	2	1.896336	0.405475		No
ravens	1119	0	4	36	24.993744	6.001412		No
english	1119	0	0	98	41.474531	21.549723		No
behaviour	1119	0	0	1	0.778374	0.415341		No
cons	1119	0	1	1	1.000000	0.000000		No

In the top left we see the status area is indicating **Finished,** as the most recent execution in this activity region has come to an end. You will notice that there are two page numbers in the navigation bar at the top, via which we can move between pages. To the left of the window is a hierarchy of the headings in the eBook. In this case I have chosen two headings each formatted as **Heading 1** and so one isn't nested within the other. The page contains the HTML text I typed, along with the summary of the data.

Next we can move to the histogram either by clicking on **Histogram** in the left navigation tree or by clicking on **2** (or **Next**) in the page navigation bar towards the top. Doing this we see:



EBook for first practical

Histogram

Here we will firstly draw a histogram of the response variable, english.

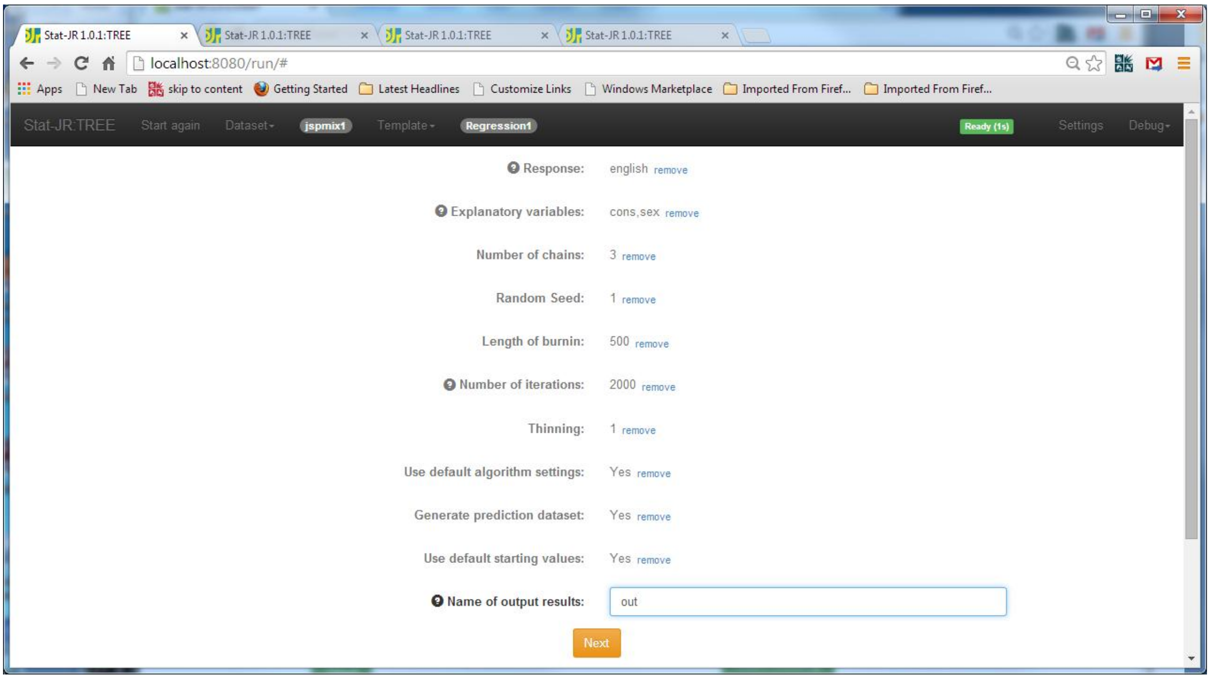
Firstly you will see the template inputs and then the histogram itself

_initial_bins	10
_initial_vals	english
bins	10
vals	english

Here we see the html text and then an inputs box which tells us which values we have chosen for this template (note these appear twice as both the initial values with a _initial_ and the current values being displayed) before finally the plot appears in its own box. This screenshot also illustrates the concept of variable page lengths, so to see the histogram fully we need to use the right hand scroll bar. As this is the whole of our eBook we cannot do much more in DEEP now so click on the **Stat-JR:DEEP** icon on the left of the top black bar to return to the main page.

Extending our E-book

We will now return to TREE where we should still have the eBook-writer screen open. In order to extend our eBook we need to actually fit the model, so click on the orange **Return to template running environment** button at the bottom of the screen. This will return us to the main TREE screen with the histogram template still selected. For now I am going to fit a first model to the data ignoring the fact that the data has two levels and so I select **Choose** from the **Template** drop-down list and scroll down and select **Regression1** and then click **Use**. We next need to set up a regression model and so I will regress **english** on **sex** as a first model (by selecting **cons** and **sex** as explanatory variables, noting the need for **cons**, an intercept column here.) In fact the inputs can be seen in the following screenshot:



Response: english remove

Explanatory variables: cons,sex remove

Number of chains: 3 remove

Random Seed: 1 remove

Length of burnin: 500 remove

Number of iterations: 2000 remove

Thinning: 1 remove

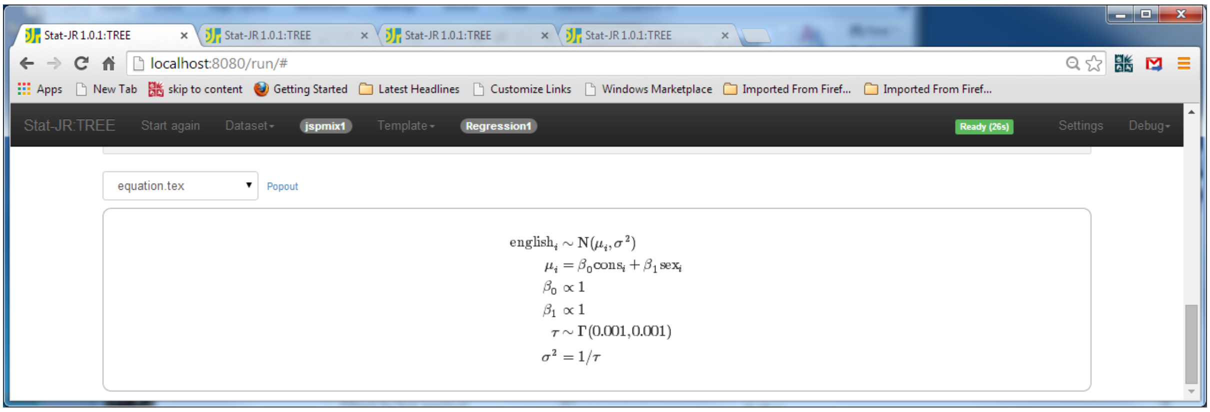
Use default algorithm settings: Yes remove

Generate prediction dataset: Yes remove

Use default starting values: Yes remove

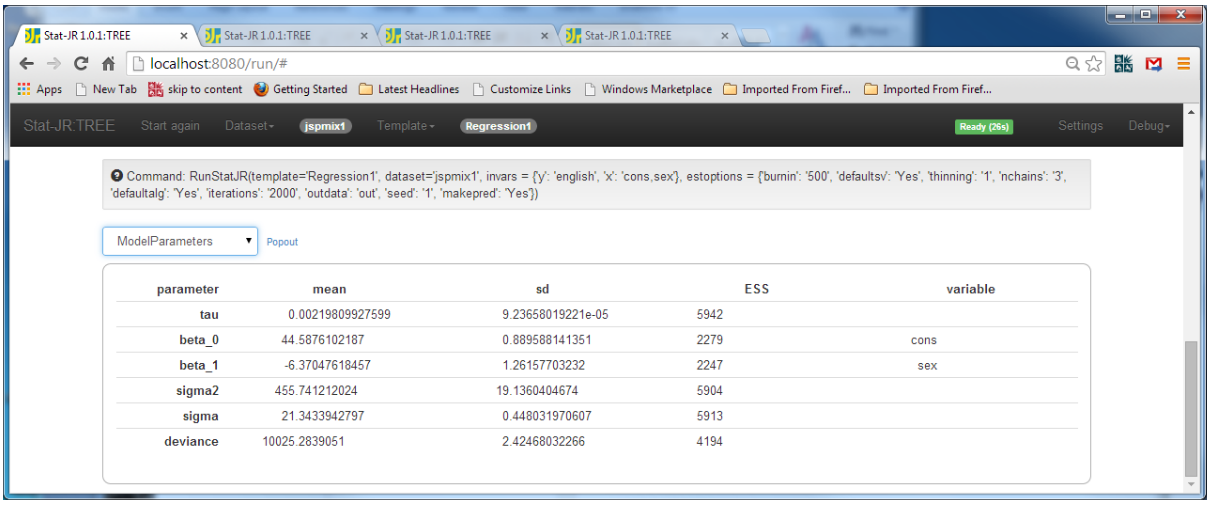
Name of output results: out

Clicking on **Next** and **Run** will fit the model using Stat-JR's built-in eSTAT MCMC engine (which is the default as it comes with the package) and this will produce lots of objects in the output list. Note that upon clicking **Run** the timer to the top-right will turn blue and say **Working** until the model has been fitted, after which it will say **Ready** again and be green. The first object that is displayed in the window is the model equation (**equation.tex**) as shown below:



$$\begin{aligned}
\text{english}_i &\sim \mathrm{N}(\mu_i, \sigma^2)\\
\mu_i &= \beta_0 \text{cons}_i + \beta_1 \text{sex}_i\\
\beta_0 &\propto 1\\
\beta_1 &\propto 1\\
\tau &\sim \Gamma(0.001, 0.001)\\
\sigma^2 &= 1/\tau
\end{aligned}$$

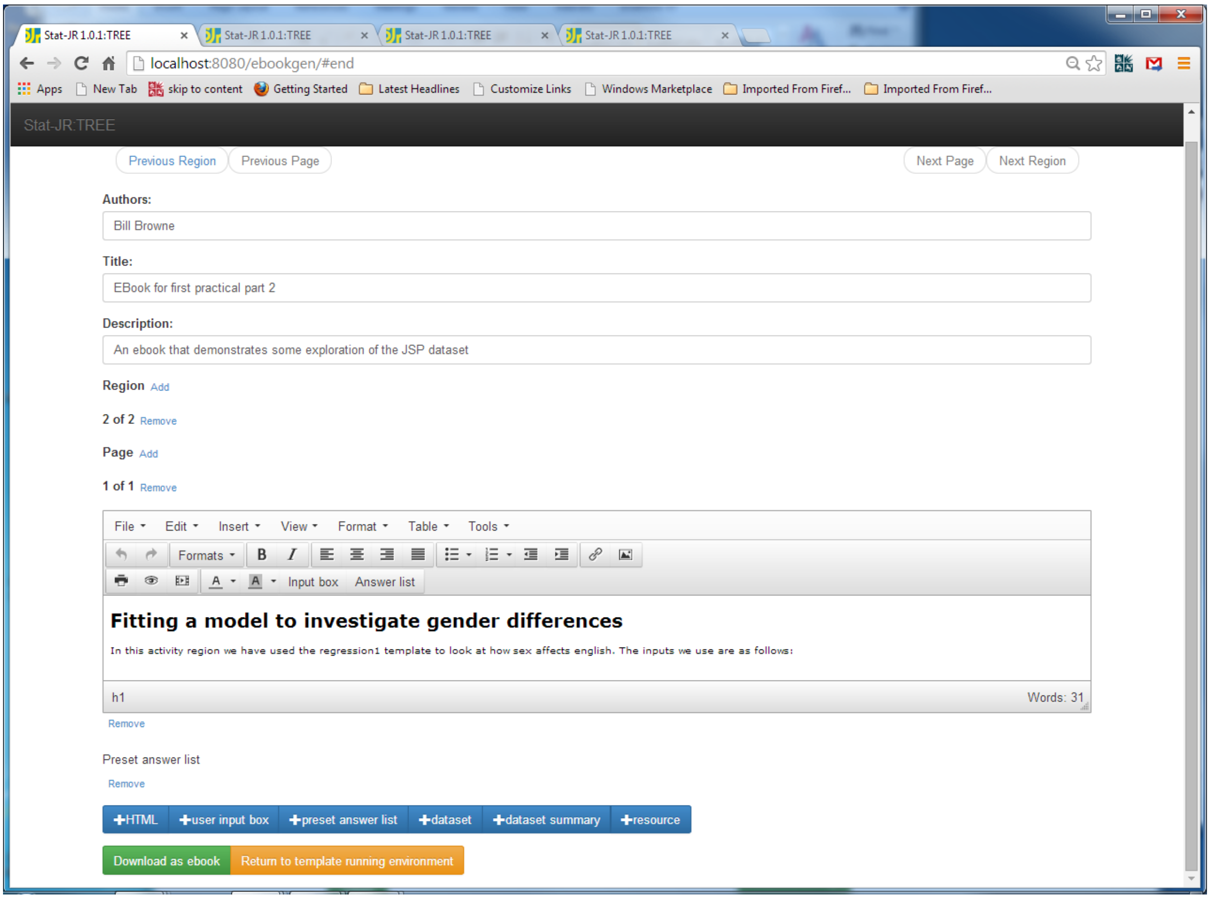
Another interesting output is the **ModelParameters** object which gives estimates for the various parameters in the model and here we see that the average score in English for sex category 0 is 44.58 whilst category 1 has an average that is -6.37 less which is significantly different:



```
Command: RunStatJR(template='Regression1', dataset='jspmix1', invars = {'y': 'english', 'x': 'cons,sex'}, estoptions = {'burnin': '500', 'defaultsv': 'Yes', 'thinning': '1', 'nchains': '3', 'defaultalg': 'Yes', 'iterations': '2000', 'outdata': 'out', 'seed': '1', 'makepred': 'Yes'})
```

parameter	mean	sd	ESS	variable
tau	0.00219809927599	9.23658019221e-05	5942	
beta_0	44.5876102187	0.889588141351	2279	cons
beta_1	-6.37047618457	1.26157703232	2247	sex
sigma2	455.741212024	19.1360404674	5904	
sigma	21.3433942797	0.448031970607	5913	
deviance	10025.2839051	2.42468032266	4194	

There are lots of other output objects we could view, including the MCMC algorithm and MCMC diagnostics plots, and we can choose any of these to include in our eBook. We will now return to the eBook-writer screen by clicking on the **Add to eBook** button. You will notice that this appears as we last left it. We will firstly change the title slightly (by adding “part 2”) so that we can differentiate this eBook from the simpler one we looked at earlier. We will also create a new activity region to contain our model-fitting; to do this click on the **Add** button next to **Region** followed by the **Add** button next to **Page** to start a first page of a second activity region. We will begin this page with some HTML, and then press the **+preset answer list** button thus:



Authors: Bill Browne

Title: EBook for first practical part 2

Description: An ebook that demonstrates some exploration of the JSP dataset

Region Add

2 of 2 Remove

Page Add

1 of 1 Remove

Fitting a model to investigate gender differences

In this activity region we have used the regression1 template to look at how sex affects english. The inputs we use are as follows:

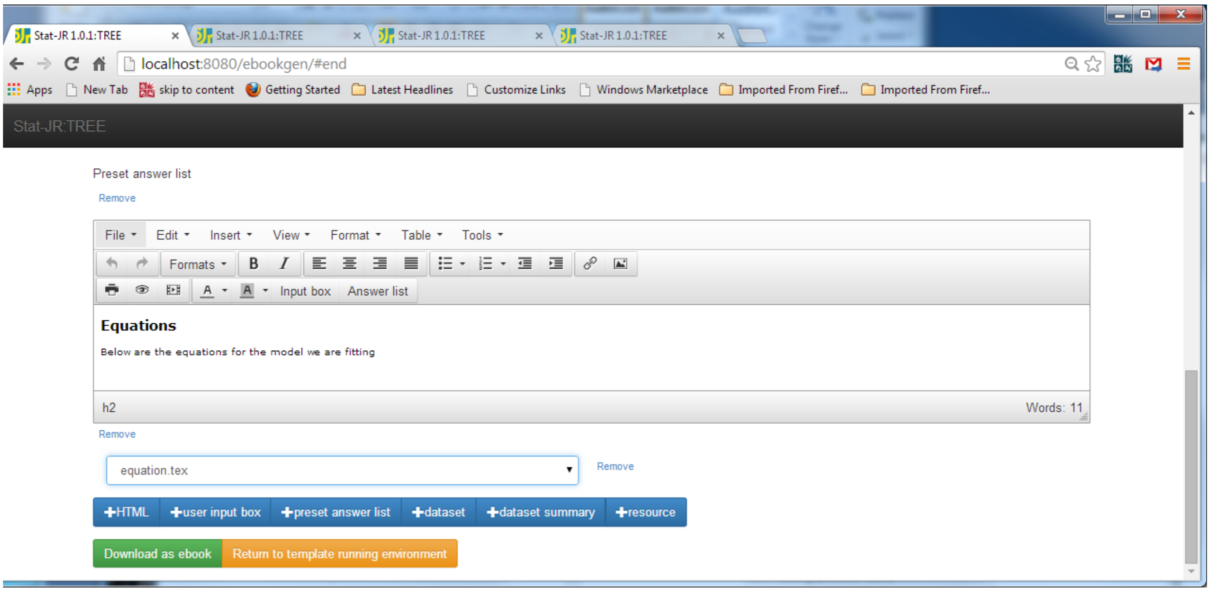
h1 Words: 31

Remove

Preset answer list

Remove

We will then add a further HTML box and the equations to the same page as follows (where for the equations we click on **+resource** and choose **equation.tex** from the **resources** list):



Preset answer list

Remove

Equations

Below are the equations for the model we are fitting

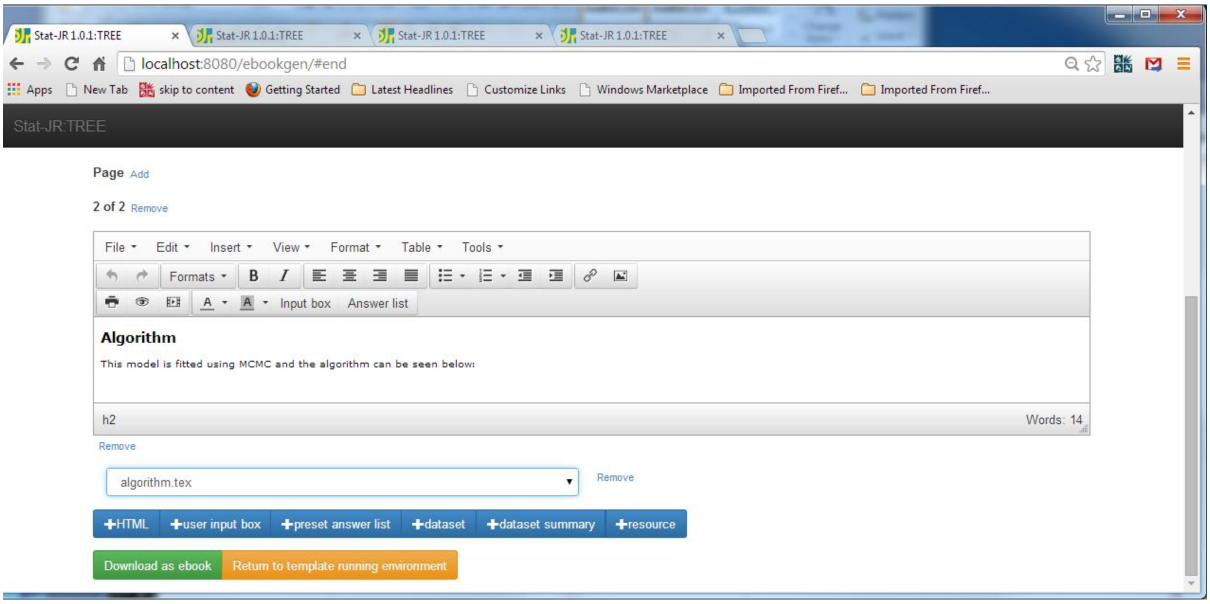
h2 Words: 11

Remove

equation.tex

Remove

Note here we have formatted the word “Equations” as **Heading 2**. We will next add a page that contains the MCMC algorithm by first clicking on **Add** next to **Page** and then adding the following this time selecting **algorithm.tex** from the resources list:



Page Add

2 of 2 Remove

Algorithm

This model is fitted using MCMC and the algorithm can be seen below:

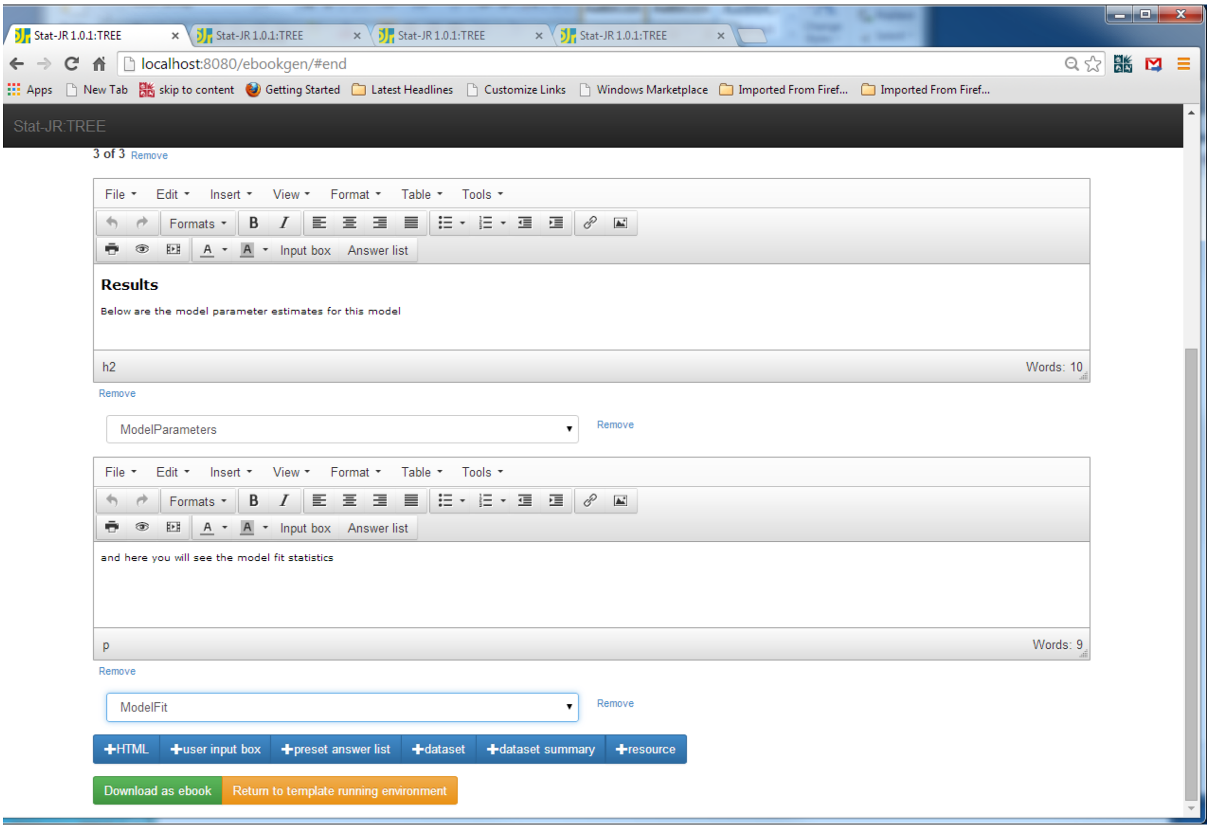
h2 Words: 14

Remove

algorithm.tex

Remove

On a third page (which we again add) we put the **ModelParameters** and the **ModelFit** objects thus:



Stat-JR 1.0.1:TREE
localhost:8080/ebookgen/#end
Apps New Tab skip to content Getting Started Latest Headlines Customize Links Windows Marketplace Imported From Firef... Imported From Firef...
Stat-JR:TREE
File Edit Insert View Format Table Tools
Formats B I
Input box Answer list
Remove
+HTML +user input box +preset answer list +dataset +dataset summary +resource
Download as ebook Return to template running environment

3 of 3

Results

Below are the model parameter estimates for this model

h2 Words: 10

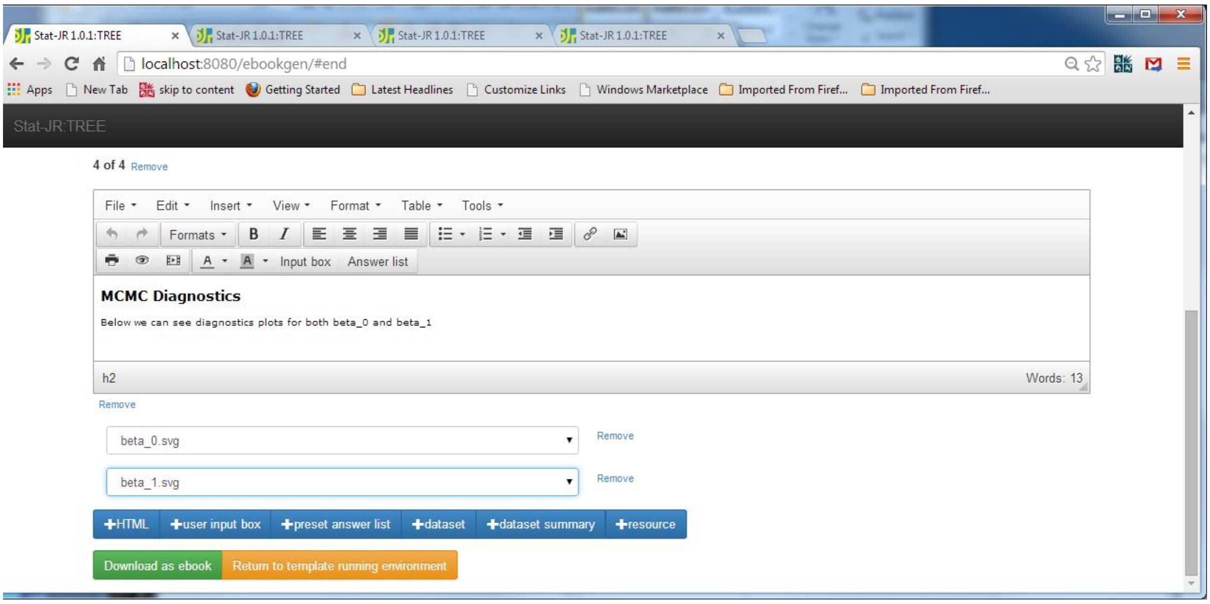
ModelParameters

and here you will see the model fit statistics

p Words: 9

ModelFit

Finally we add a couple of MCMC diagnostics plots on a fourth page:



Stat-JR 1.0.1:TREE
localhost:8080/ebookgen/#end
Apps New Tab skip to content Getting Started Latest Headlines Customize Links Windows Marketplace Imported From Firef... Imported From Firef...
Stat-JR:TREE
File Edit Insert View Format Table Tools
Formats B I
Input box Answer list
Remove
+HTML +user input box +preset answer list +dataset +dataset summary +resource
Download as ebook Return to template running environment

4 of 4

MCMC Diagnostics

Below we can see diagnostics plots for both beta_0 and beta_1

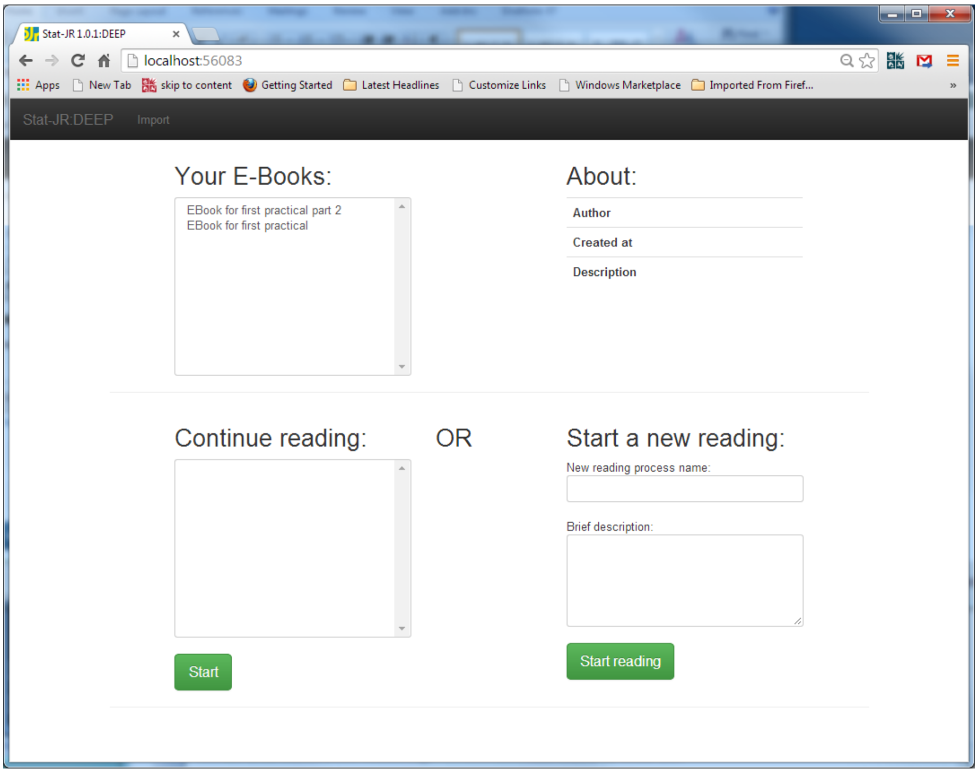
h2 Words: 13

beta_0.svg

beta_1.svg

We have now added four pages to this activity region (to make our eBook six pages long) and although we have selected some of the objects produced by this template execution we have by no means selected all of them. We have, however, completed what we want for this eBook and so we can now download it and this time save it under a new name, in my case **jsp2.zip**.

Returning to the DEEP interface we can now click on **Import** and load up this second eBook to our list thus:



Your E-Books:

EBook for first practical part 2
EBook for first practical

About:

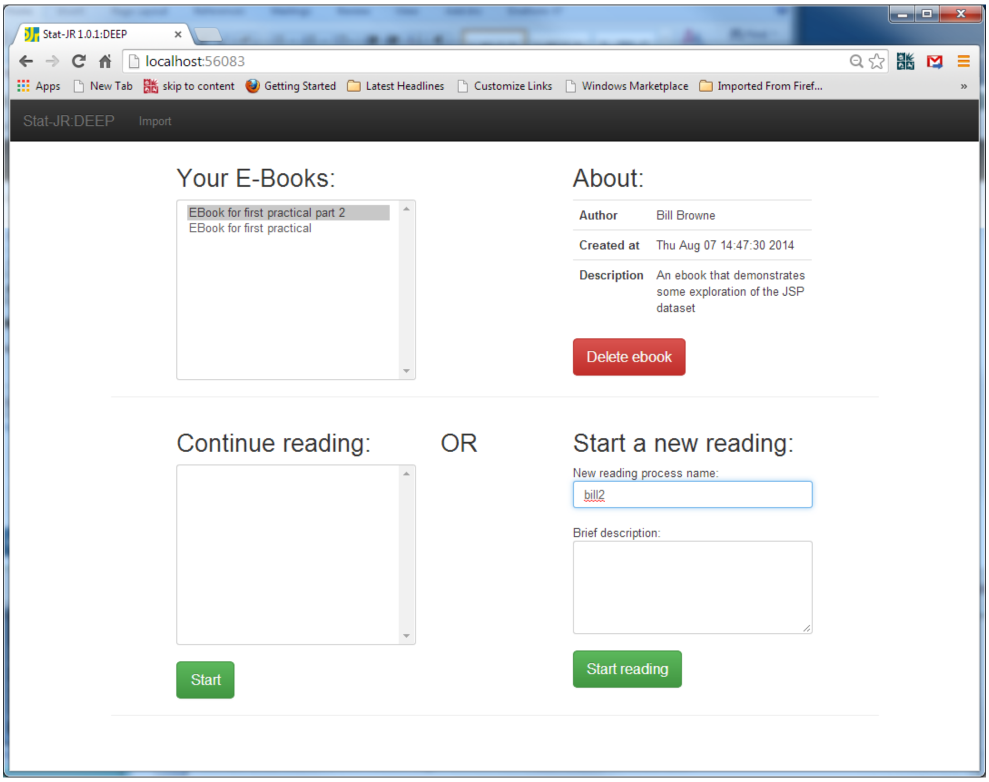
Author	
Created at	
Description	

Continue reading: OR Start a new reading:

New reading process name:

Brief description:

You can see above why it is important to change the name. Next, click on the new eBook and give it a reading process name:



Your E-Books:

EBook for first practical part 2
EBook for first practical

About:

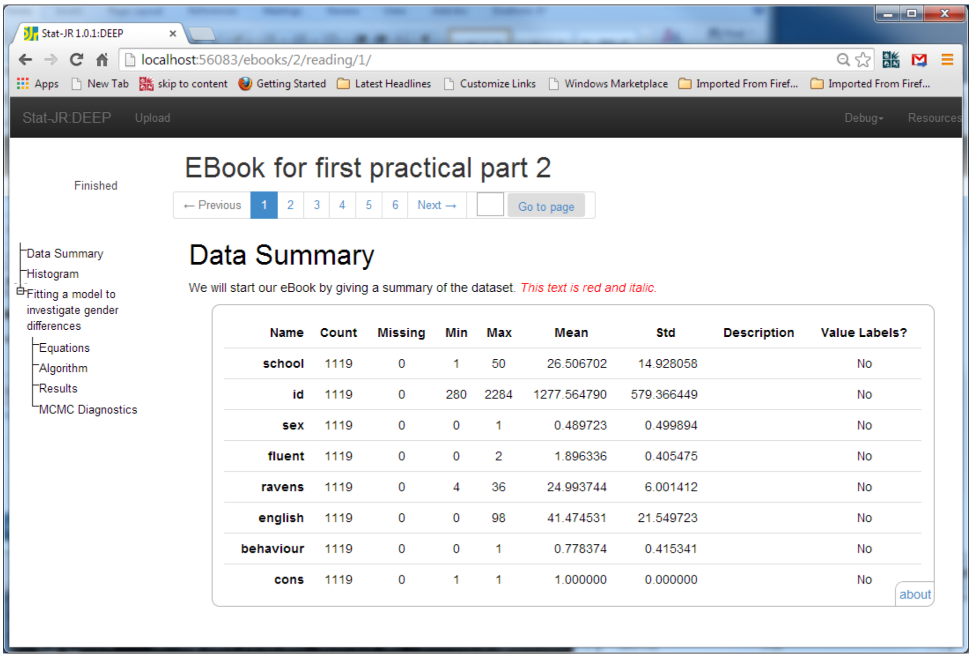
Author	Bill Browne
Created at	Thu Aug 07 14:47:30 2014
Description	An ebook that demonstrates some exploration of the JSP dataset

Continue reading: OR Start a new reading:

New reading process name: bill2

Brief description:

We can now click on **Start reading** and go through the six pages in turn thus:


Stat-JR 1.0.1:DEEP
localhost:56083/ebooks/2/reading/1/
Apps New Tab skip to content Getting Started Latest Headlines Customize Links Windows Marketplace Imported From Firef... Imported From Firef...
Stat-JR:DEEP Upload Debug Resources
Finished
← Previous 1 2 3 4 5 6 Next → Go to page
Data Summary
Histogram
Fitting a model to investigate gender differences
Equations
Algorithm
Results
MCMC Diagnostics
about

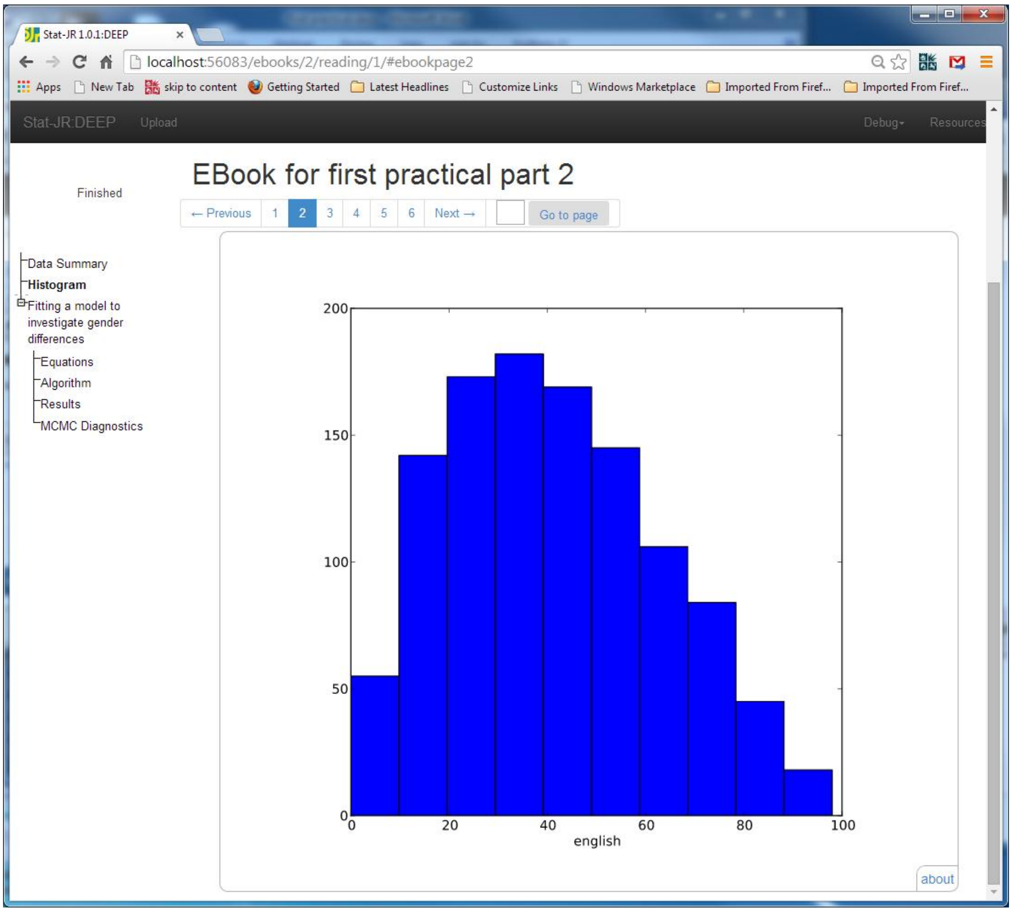
EBook for first practical part 2

Data Summary

We will start our eBook by giving a summary of the dataset. *This text is red and italic.*

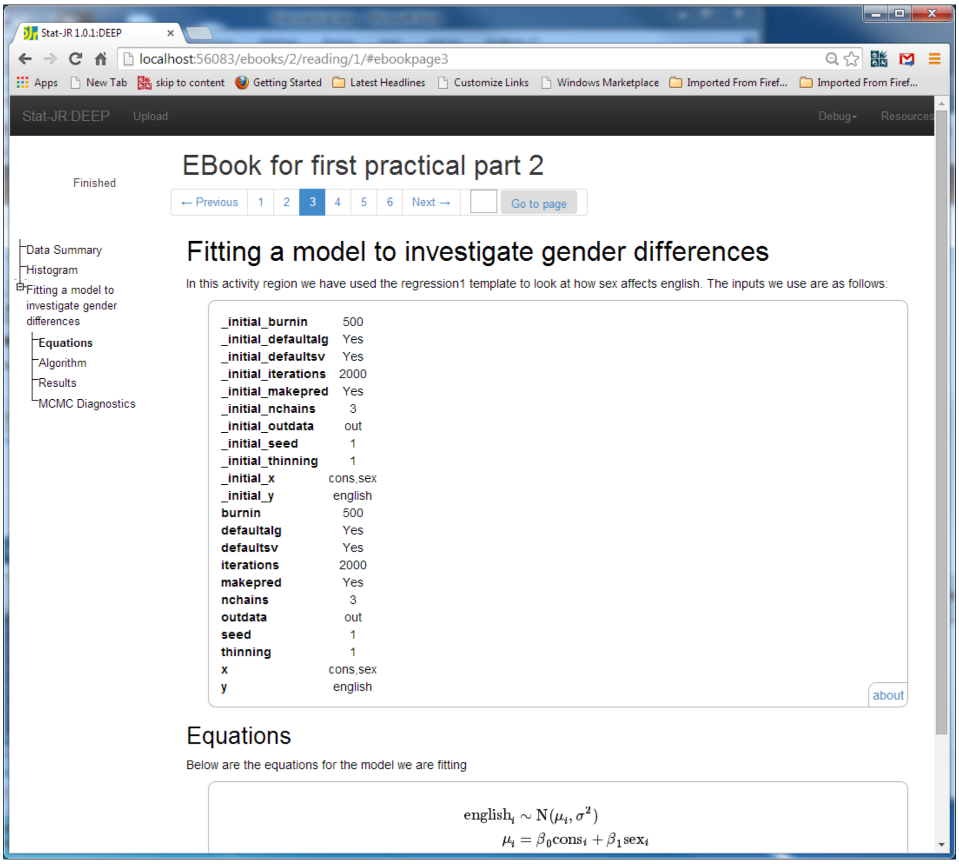
Name	Count	Missing	Min	Max	Mean	Std	Description	Value Labels?
school	1119	0	1	50	26.506702	14.928058		No
id	1119	0	280	2284	1277.564790	579.366449		No
sex	1119	0	0	1	0.489723	0.499894		No
fluent	1119	0	0	2	1.896336	0.405475		No
ravens	1119	0	4	36	24.993744	6.001412		No
english	1119	0	0	98	41.474531	21.549723		No
behaviour	1119	0	0	1	0.778374	0.415341		No
cons	1119	0	1	1	1.000000	0.000000		No

The first page of the eBook is as we have seen before but note the additional page numbers listed at the top of the screen and the hierarchical heading list to the left. Next, looking at page 2 (having scrolled down the screen), we see:


Stat-JR 1.0.1:DEEP
localhost:56083/ebooks/2/reading/1/#ebookpage2
Apps New Tab skip to content Getting Started Latest Headlines Customize Links Windows Marketplace Imported From Firef... Imported From Firef...
Stat-JR:DEEP Upload Debug Resources
Finished
← Previous 1 2 3 4 5 6 Next → Go to page
Data Summary
Histogram
Fitting a model to investigate gender differences
Equations
Algorithm
Results
MCMC Diagnostics
200
150
100
50
0
0
20
40
60
80
100
english
about

EBook for first practical part 2

Clicking on the next page moves us onto the next activity region and the model now gets fitted:



EBook for first practical part 2

Fitting a model to investigate gender differences

In this activity region we have used the regression1 template to look at how sex affects english. The inputs we use are as follows:

_initial_burnin	500
_initial_defaultalg	Yes
_initial_defaultsv	Yes
_initial_iterations	2000
_initial_makepred	Yes
_initial_nchains	3
_initial_outdata	out
_initial_seed	1
_initial_thinning	1
_initial_x	cons,sex
_initial_y	english
burnin	500
defaultalg	Yes
defaultsv	Yes
iterations	2000
makepred	Yes
nchains	3
outdata	out
seed	1
thinning	1
x	cons,sex
y	english

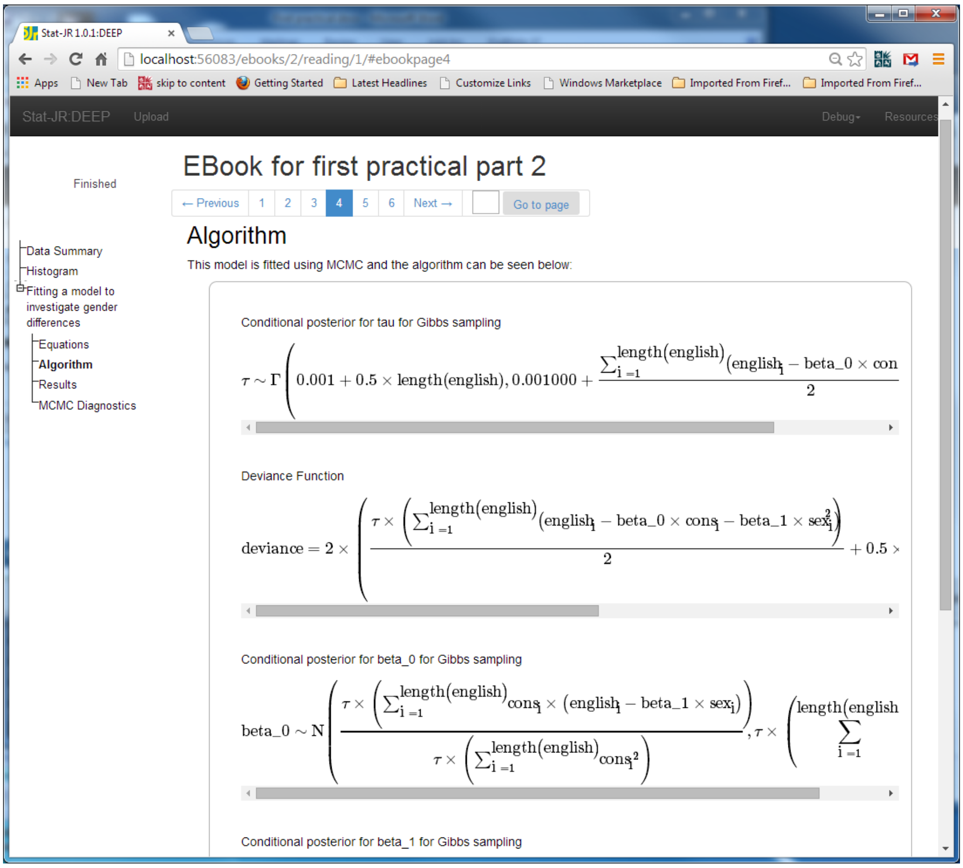
Equations

Below are the equations for the model we are fitting

$$\mathrm{english}_i \sim \mathrm{N}(\mu_i, \sigma^2)$$

$$\mu_i = \beta_0 \mathrm{cons}_i + \beta_1 \mathrm{sex}_i$$

Again we would need to scroll down to see the equations in their entirety. On page four we find the algorithm steps:



EBook for first practical part 2

Algorithm

This model is fitted using MCMC and the algorithm can be seen below:

Conditional posterior for tau for Gibbs sampling

$$\tau \sim \Gamma\Bigg(0.001 + 0.5 \times \mathrm{length(english)}, 0.001000 + \frac{\sum_{i=1}^{\mathrm{length(english)}}(\mathrm{english}_i - \mathrm{beta\_0} \times \mathrm{con}}{2}$$

Deviance Function

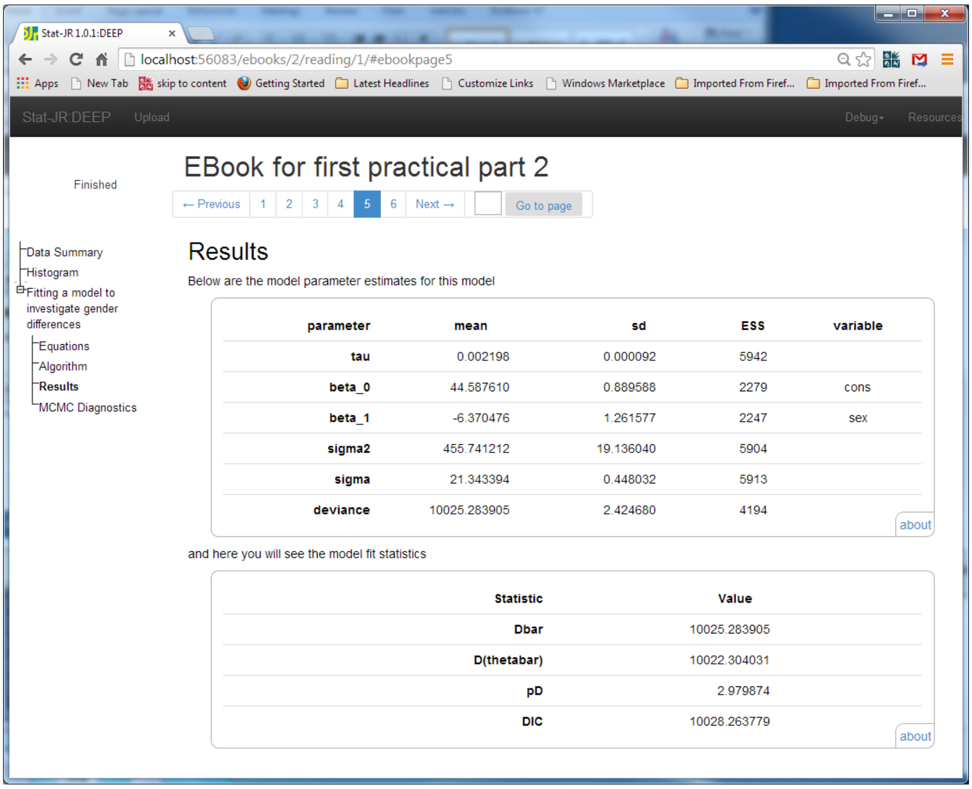
$$\mathrm{deviance} = 2 \times \Bigg(\frac{\tau \times \left(\sum_{i=1}^{\mathrm{length(english)}}(\mathrm{english}_i - \mathrm{beta\_0} \times \mathrm{cons}_i - \mathrm{beta\_1} \times \mathrm{sex}_i)^2\right)}{2} + 0.5 \times$$

Conditional posterior for beta_0 for Gibbs sampling

$$\mathrm{beta\_0} \sim \mathrm{N}\Bigg(\frac{\tau \times \left(\sum_{i=1}^{\mathrm{length(english)}}\mathrm{cons}_i \times (\mathrm{english}_i - \mathrm{beta\_1} \times \mathrm{sex}_i)\right)}{\tau \times \left(\sum_{i=1}^{\mathrm{length(english)}}\mathrm{cons}_i^2\right)}, \tau \times \Bigg(\sum_{i=1}^{\mathrm{length(english}}$$

Conditional posterior for beta_1 for Gibbs sampling

On page five we see the results and the fit statistics:



EBook for first practical part 2

Results

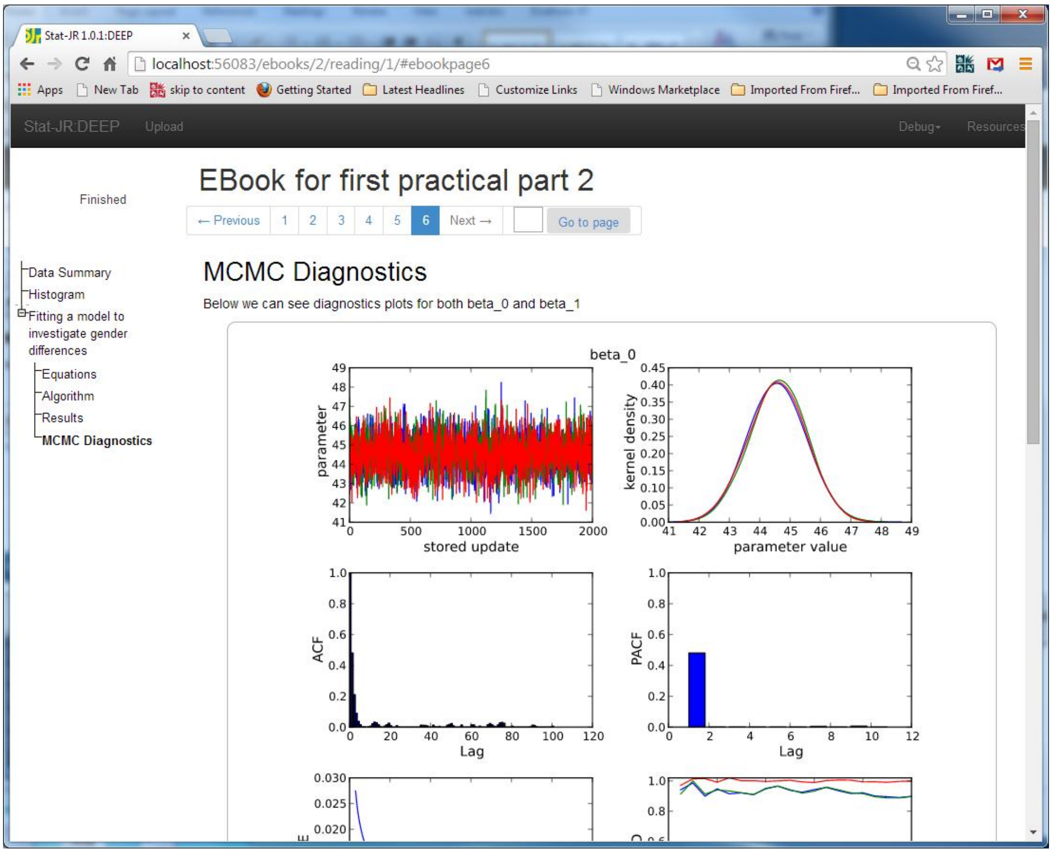
Below are the model parameter estimates for this model

parameter	mean	sd	ESS	variable
tau	0.002198	0.000092	5942	
beta_0	44.587610	0.889588	2279	cons
beta_1	-6.370476	1.261577	2247	sex
sigma2	455.741212	19.136040	5904	
sigma	21.343394	0.448032	5913	
deviance	10025.283905	2.424680	4194	

and here you will see the model fit statistics

Statistic	Value
Dbar	10025.283905
D(thetabar)	10022.304031
pD	2.979874
DIC	10028.263779

And finally the MCMC diagnostics plots on the last page:

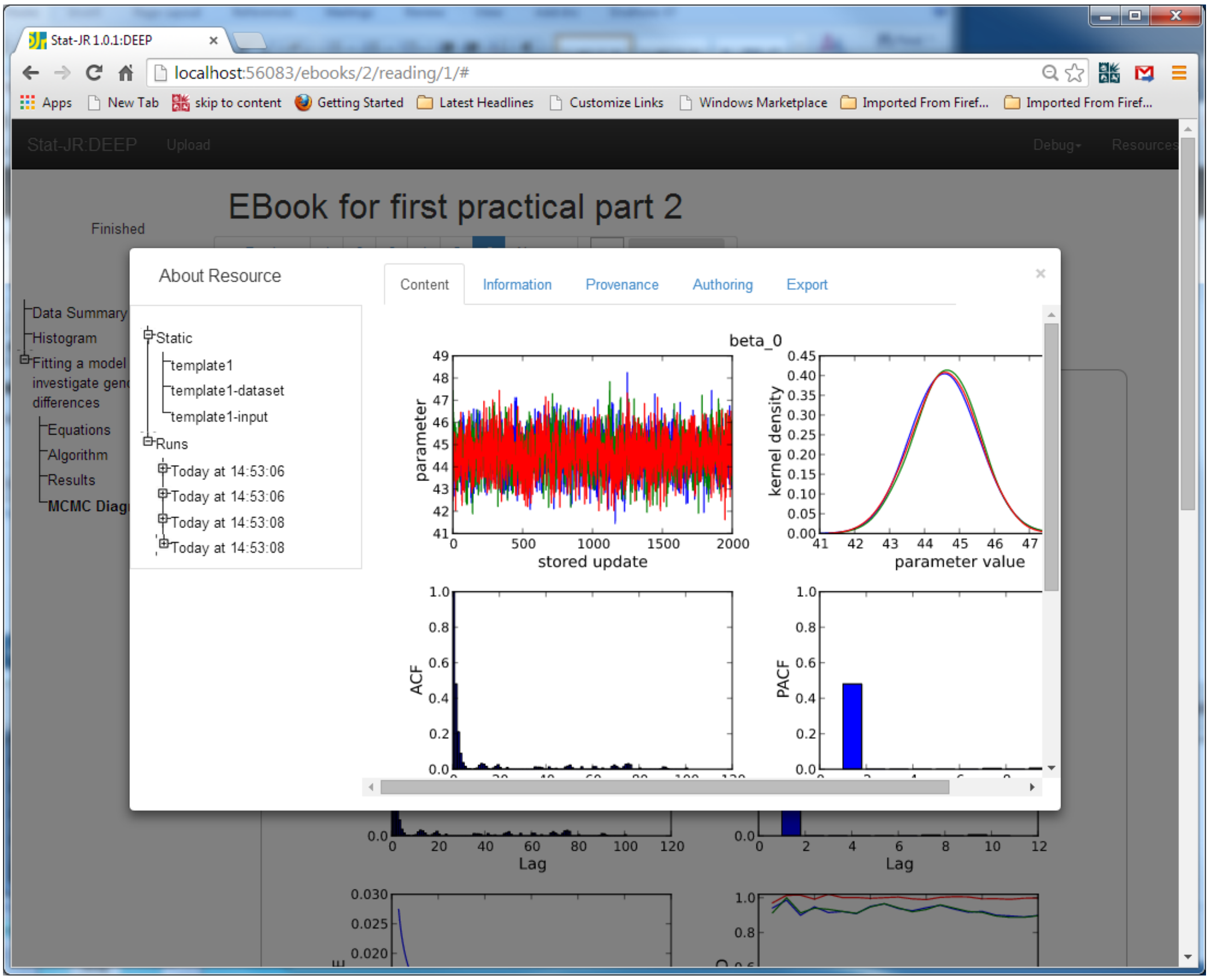


EBook for first practical part 2

MCMC Diagnostics

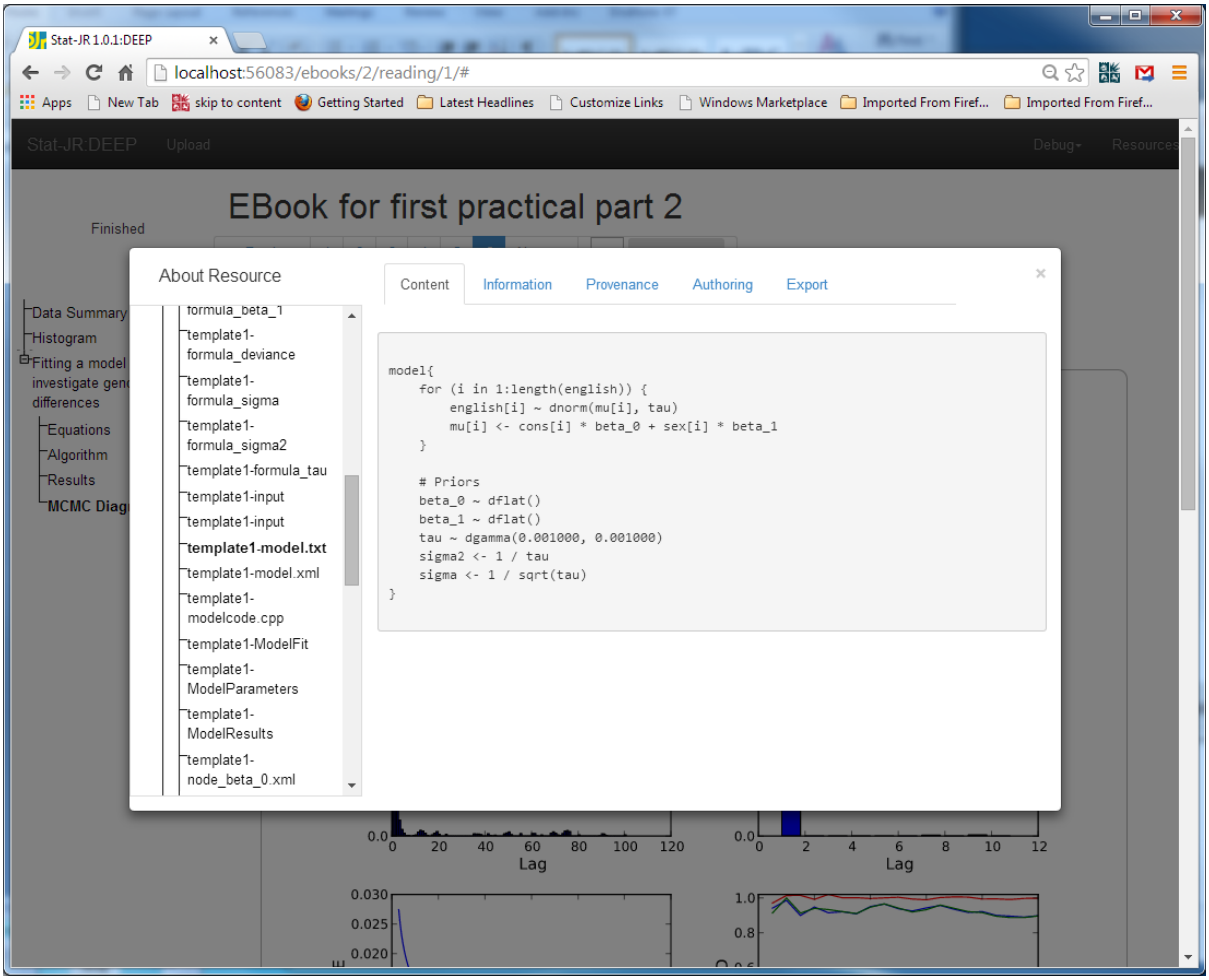
Below we can see diagnostics plots for both beta_0 and beta_1

As mentioned earlier, not all objects created by Stat-JR need appear in the front-end of an eBook, although they are all accessible behind-the-scenes in DEEP. For example, clicking on the **about** button that appears in the bottom-right corner of the **beta_0** graph takes you to the **About Resource**[2] box, as follows:



To the left you will see a list of objects organised under the headings **Static** and **Runs. Static** groups things that do not change in the eBook for example the template code and the dataset while the list of **Runs** is added to upon each execution performed while viewing the eBook. Under **Runs** you'll find the dynamic executions the system has performed, arranged by the time they were completed. Clicking on the **+** sign next to one of these runs opens a list of all the objects constructed as part of that execution. For example, in my case the second run at 14:53:06 produced a large number of output objects and from this I can view, for instance, the model code that Stat-JR uses for the model fitting (**template1-model.txt**) even though it doesn't appear in the front-end of the eBook, as shown below:

[2] Alternatively if you didn't want to jump straight to a specific resource you can click on **Resources** to the right of the black bar at the top.



```
model{
    for (i in 1:length(english)) {
        english[i] ~ dnorm(mu[i], tau)
        mu[i] <- cons[i] * beta_0 + sex[i] * beta_1
    }

    # Priors
    beta_0 ~ dflat()
    beta_1 ~ dflat()
    tau ~ dgamma(0.001000, 0.001000)
    sigma2 <- 1 / tau
    sigma <- 1 / sqrt(tau)
}
```

This ends our first practical. If you had sped through it then by all means investigate using further templates and including them in your eBook. By the end of the practical you should be comfortable using the TREE interface to perform analysis in Stat-JR and the eBook-writer to create a logfile-style eBook of the analysis performed in TREE within DEEP.

In the next practical we will make an eBook that is somewhat more interactive.