



Bookmarks

- ▶ Important Pre-Course Survey
- ▶ Contact Us
- ▶ How To Navigate the Course
- ▶ Discussion Board
- ▶ Office Hours
- ▶ Week 1: Introduction to Data
- ▶ Week 2: Univariate Descriptive Statistics
- ▶ Week 3: Bivariate Distributions
- ▶ Week 4: Bivariate Distributions (Categorical Data)
- ▶ Week 5: Linear Functions
- ▼ **Week 6: Exponential and Logistic**

Week 6: Exponential and Logistic Function Models &gt; Pre-Lab &gt; Examine the Data



Bookmark

Reflect on the Question

Analyze the Data

Draw Conclusions

## Lab 6: Worldwide Trends in Internet Usage



The World Bank is a data collection of information on all the world's countries. Data is collected by country, and include items such as total population, CO<sup>2</sup> emissions, and the number of mobile device subscriptions. We will examine some of the trends in this dataset and interpret the parameters of the fitted models to best describe the change over time.

### Primary Research Question

What model best describes the first decade of internet usage (1990-1999) in the United States? Which model is a better long-term fit?


(3/3 points)

### Check the Data


Let's begin by examining our data in R.

## Function Models

### Readings


Reading Check due Mar 15, 2016 at 18:00 UTC 

### Lecture Videos

Comprehension Check due Mar 15, 2016 at 18:00 UTC 

### R Tutorial Videos


### Pre-Lab

Pre-Lab due Mar 15, 2016 at 18:00 UTC 

### Lab

Lab due Mar 15, 2016 at 18:00 UTC 

### Problem Set

Problem Set due Mar 15, 2016 at 18:00 UTC 

1. Open RStudio. Make sure you've installed the SDSFoundations package.
2. Type `library(SDSFoundations)` This will automatically load the data for the labs.
3. Type `world <- WorldBankData` This will assign the data to your Workspace.
4. Look at the spreadsheet view of the data to answer the following questions.

**Alternatively**, you can use follow the steps in the "Importing a Data Frame" R tutorial video, and use the WorldBankData.csv file. (Right-click and "Save As.") Make sure to **name** the dataframe "world" when importing.

1. Open RStudio.
2. Click on "Import Dataset" button at the top of the workspace window. Choose "from text file."
3. Click on the location of the WorldBankData.csv file you just downloaded.
4. Click on the WorldBankData.csv file. Then, click Upload.
5. Look at the spreadsheet view of the data to answer the following questions.

1a) What is the first "Low Income" country in the dataset?



Answer: Afghanistan

1b) What was the rural population of Aruba in 1970? (*Report without commas*)



Answer: 29164

1c) When was the first year Australia had data on the number of mobile device subscriptions? (Subscriptions more than 0)



Answer: 1987

[Click here for a video explanation of how to answer this question.](#)

*You have used 1 of 1 submissions*

(4/4 points)

## Check the Variables of Interest

Let's find the variables we need to answer the question.

2a) Which variable tells us the *number* of internet users in a specific year? The variable name in the dataset is:

internet.users ▾



Answer: internet.users

2b) What type of variable is this?

quantitative ▾



Answer: quantitative

2c) Which variable tells us *when* the number of internet users was recorded? The variable name in the dataset is:

year ▾



Answer: year

2d) What type of variable is this?

quantitative ▾



Answer: quantitative

[Click here for a video explanation of how to answer this question.](#)

*You have used 1 of 1 submissions*

(2/2 points)

## Reflect on the Method

*Which method should we be using for this analysis and why?*

3a) What **statistic** helps us determine how well a particular model fits the data?

☐ Variance

☒ R-squared

☐ Mean

3b) In this lab, we will calculate **residuals** after fitting both an exponential and a logistic model to a set of data. What is a **residual**?

- ☒ A residual is the difference between a predicted value and the actual observed value. ✓
- ☐ A residual is the average distance of a data point to the line of best fit.
- ☐ A residual is a statistic that tells you how much variation is explained by the model.

[Click here for a video explanation of how to answer this question.](#)

*You have used 1 of 1 submissions*

© All Rights Reserved



© edX Inc. All rights reserved except where noted. EdX, Open edX and the edX and Open EdX logos are registered trademarks or trademarks of edX Inc.

POWERED BY  
OPENedX

