

# Universal comparison of school education in RStudio

**Mahesh Mardolkar**

Research Scholar  
Department of Computer Science & Engg  
Annamalai University  
Tamil Nadu

**N. Kumaran**

Assistant Professor  
Department of Computer Science & Engg  
Annamalai University  
Tamil Nadu

**Abstract:** *R studio is an Integrated Development Environment(IDE) which is free and is open source software developed for R programming. Through user created packages the different capabilities of R are extended. These packages are developed in R and other languages like JAVA, C, C++ and Fortran. ggplot2 package- is used for data visualization.*

*The number of universities established stands at 789 with over 37, 204 colleges and another 11, 443 stand alone institution in India Domestic activities, marriage and lack of interest is the key reasons for female students to dropout, on the other hand lack of interest, economic activities and financial constraints are the reasons for male students to dropout.*

**Keywords:** *Dropout, R programming, ggplot2*

## Introduction

The number of universities established stands at 789 with over 37, 204 colleges and another 11, 443 stand alone institution in India, also the Eight All India School Education Survey(AISES) has put a total of 13, 06,992 schools and enrolment figure as 22,67,19,283 number of students in the school, the problem is in continuing studies at school or college for many students is difficult due to reason like financial problem, domestic problems at home, unable to cope up with studies, language or medium of education, and many other problems like this which make it difficult to continue their education, under such circumstances the student drops further studies. The comparative study of school education is shown in RStudio by implementing ggplot2 package for generating the graph to present the education scenario.

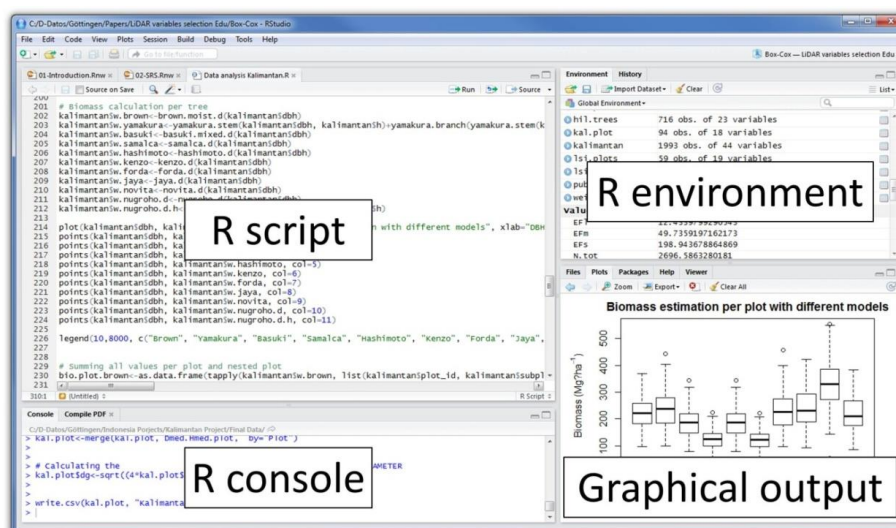


Fig 1 : RStudio Desktop Interface

## RStudio

Is an Integrated Development Environment (IDE) which is free and is open source software developed for R programming. R is a programming language mainly used for statistical computing and for graphics. J J Allair the creator of ColdFusion programming language also founded RStudio, two editions of RStudio are available for user, RStudio desktop where user can run local desktop applications and RStudio Server, which uses a web browser to access RStudio an remote linux server. Distributions of RStudio is available for Linux, Windows and Mac OS, commercial and open source editions of RStudio is available for the users. RStudio is created using C++ programming language, RStudio also used Qt framework for its graphical user interface. The first public beta version was announced in February 2011 on November 2016 version 1.0 was released and on 9 October 2017 version 1.1 was released.

## Packages

Through user created packages the different capabilities of R are extended. These packages are developed in R and other languages like JAVA, C, C++ and Fortran. Packages allow specialized statistical techniques to be incorporated easily like graphical devices (ggplot2), Import/ Export capabilities, reporting tools (knitr, Sweave), etc. R installation includes core set of packages and additional packages of more than 7,801 available at Comprehensive R Archive Network(CRAN), Bioconductor, Omegahat, GitHub and other repositories.

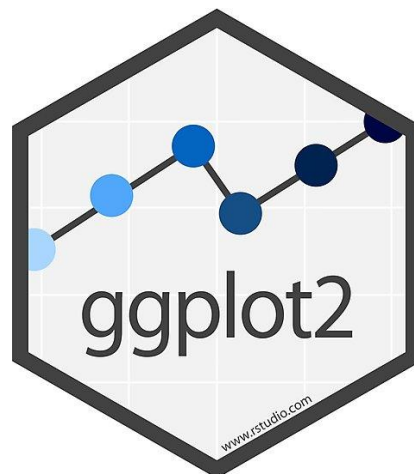


Fig 2:ggplot2 package logo

ggplot2 package- is used for data visualization. Hadley Wickham created ggplot package in 2005. Leland Wilkinson's grammar of graphics is implemented in ggplot. The basic graphics in R can be replaced by ggplot2 which contains print display and number of defaults for web. ggplot2 is one of the most popular package of R license under GNU GPL.

## Updates

ggplot2 version 0.9 was released on 2 March 2012, after long time of maintenance mode operations ggplot2.0.0 was released on 21 December 2015. Comparison with other packages and base graphics – with the high level of abstraction ggplot2 allows user to alter, add or remove components in a plot. The cost of this abstraction is ggplot2 is slower than lattice graphics, more complex plotting can be done in ggplot2.

## Education in Schools

The learning outcome in school was recognized in union budget speech. Right to Education(RTE) Act was introduced and an amendment was passed to prevent detention after class V and class VIII after taking test, and a retest and also after remedial education.

Gross Enrolment Ratio(GER) is student enrolment with corresponding eligibility age group, at elementary level class I-VIII universal enrolment has been achieved. GER of 114% in 2008-09 is reduced to 100% in 2014-15 for class I-V. The class enrolment in 2008-09 is above 100%, which includes younger than six years and older than 10 years, it was 100% in 2014-15 since class composition was more appropriate.

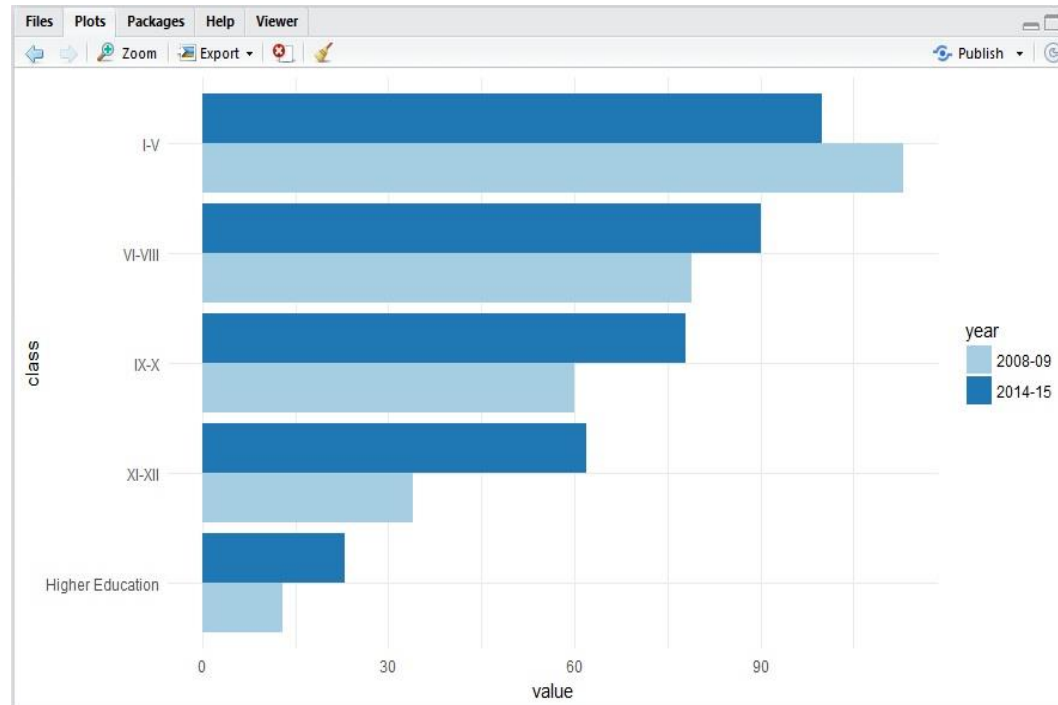


Chart 1: Gross Enrolment Ratio (GER)

```
> a <- read.csv("d:/backup/paper/class.csv")
> ggplot(a, aes(x=class,y=value, fill=year)) +geom_bar(stat="identity",
position=position_dodge())+scale_fill_brewer(palette="Paired")+theme_minimal()
+ scale_x_discrete(limits=c("Higher Education", "XI-XII", "IX-X", "VI-VIII", "I-
V")) + coord_flip()
>
```

### R Script 1: Gross Enrolment Ratio (GER)

In comparison with developed countries India's enrolment rate in primary education class I-V seems to be better and fall behind after class VI. India's enrolment rate stands at 23% against 87%, 57% and 39% in US, UK and China respectively in higher education.

The drop out ratio in 2014-15 for secondary level class IX to X is at 17% as compared to 4% in primary class I to V, 4% in upper primary class VI to VIII and 2% in upper secondary class XI to XII. The transition rate from class X to class XI is at 69% since many students are held back or have dropped out of school, until the completion of elementary education a child cannot be expelled or detained under the RTE Act.

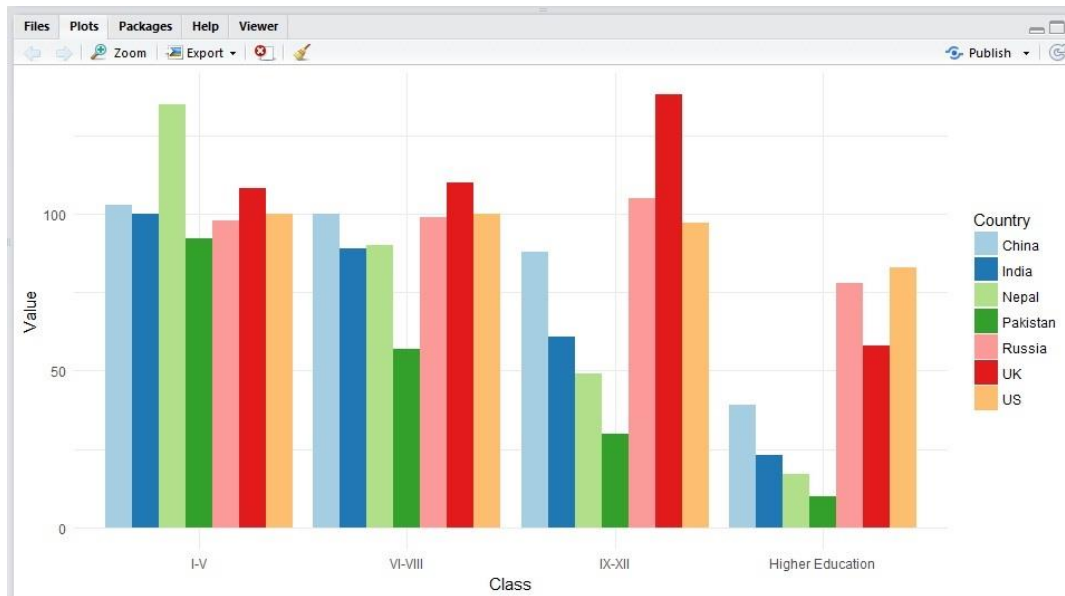


Chart 2: Countrywise comparison of enrolment

```
> a <- read.csv("d:/backup/paper/country.csv")
ggplot(a, aes(x=Class,y=Value, fill=Country)) +geom_bar(stat="identity",
position=position_dodge())+scale_fill_brewer(palette="Paired")+theme_minimal()
+ scale_x_discrete(limits=c("I-V", "VI-VIII", "IX-XII", "Higher Education"))
```

R Script 2: Countrywise enrolment

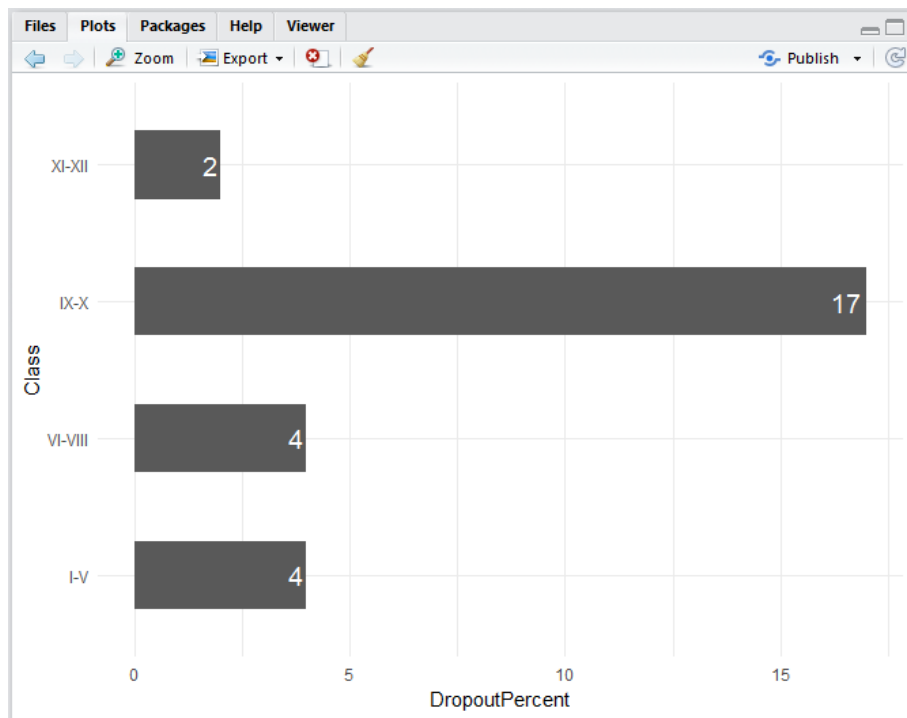


Chart 3: Classwise Dropout Ratio

```

> a <- read.csv("d:/backup/paper/dropout.csv")
> ggplot(a, aes(x=Class, y=DropoutPercent)) +geom_bar(stat="identity",
position=position_dodge(), width=0.5)+scale_fill_brewer(palette="Paired")+geom_text
(aes(label=DropoutPercent), vjust=0.5, hjust=1.2, color="white", position =
position_dodge(0.5), size=5)+ theme_minimal() + scale_x_discrete(limits=c("I-
V", "VI-VIII", "IX-X", "XI-XII")) + coord_flip()

```

#### R Script 3: Classwise Dropout Ratio

Domestic activities, marriage and lack of interest is the key reasons for female students to dropout, on the other hand lack of interest, economic activities and financial constraints are the reasons for male students to dropout.

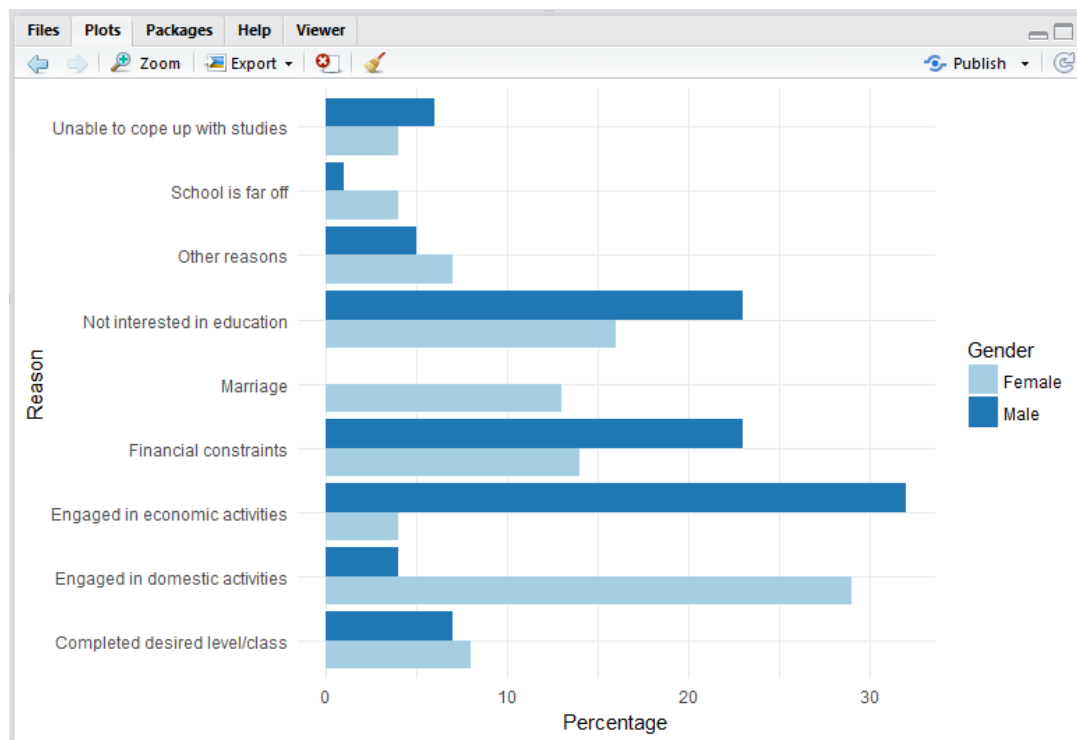


Chart 4: Reasons for Dropout

```

> a <-read.csv("d:/backup/paper/Reasons.csv")
> ggplot(a, aes(x=Reason,y=Percentage, fill=Gender)) +geom_bar(stat="identity",
position=position_dodge())+scale_fill_brewer(palette="Paired")+theme_minimal() +
coord_flip()
>

```

#### R Script 4: Reasons for Dropout

### Conclusion

All India Higher Education Survey(AIHES) an initiative of Union Human Resource Development(HRD) has registered an increase in India's Gross Enrolment Ratio(GER) from 24.5% in 2015-16 to 25.2 in 2016-17. The number of students enrolled in undergraduate and postgraduate level is measured and expressed as percentage known as GER. India aims to attain 30% GER by 2020. Enrolment in higher classes in schools has improved, India's

enrolment levels are much worse than those of developed nations. Dropout rates are highest for class IX and X, male students dropout to work while female students are occupied with domestic activities.

### References

1. Daniel T. Larose, Chantal D. Larose Discovering knowledge in Data Second Edition John Wiley
2. Trever Hstie, Robert Tibshirani, Jerome Friedman, The Elements of Statistical Learning First Edition 2001 Springer
3. Richard Cotton, Learning R O' REILLY
4. Peter Dalgaard, Introductory Statistics with R Second Edition Springer
5. Mahesh Mardolkar, N Kumaran. School Dropout Analysis with R programming charts. International Journal of Research(IJR), Vol-05, Issue-04, February 2018 ISSN:2348-6848.
6. <https://www.livemint.com/Education/k1ANVHwheaCFWCupY3jkFP/Trends-in-school-enrolment-and-dropout-levels.html>
7. [http://wiki.awf.forst.uni-goettingen.de/wiki/index.php/Installation\\_and\\_Interface\\_of\\_R](http://wiki.awf.forst.uni-goettingen.de/wiki/index.php/Installation_and_Interface_of_R)
8. <https://www.redbubble.com/people/rstudio-inc/works/31171461-ggplot2-hex-logo?p=canvas-print>
9. <https://en.wikipedia.org/wiki/RStudio>
10. <https://www.r-graph-gallery.com/portfolio/ggplot2-package/>
11. <https://ggplot2.tidyverse.org/>
12. <https://indianexpress.com/article/education/indias-gross-enrolment-ratio-in-higher-education-up-by-0-7-5012579/>