

Session 6 - Tableau for Data Science

Case study eight

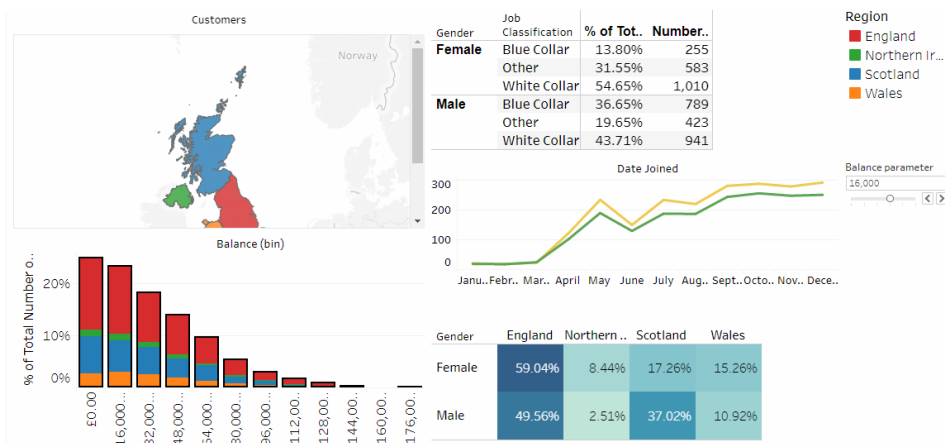
Scenario: Imagine you are working as product analyst at a bank and you are tasked with analysis to identify customer segments, for targeting them with relevant financial products. Perform descriptive data analysis and summarize using interactive dashboards

Dataset (press CTRL+ CLICK to open the link): [Download](#)

Dataset description: Segmentation of Bank Customers in United Kingdom

Data source: Superdatascience web

A sample representation of the customer segmentation dashboard



1. Load the data set and analyze the nature of variables in the dataset
2. Transform “Regions” dimension (i.e. variable) into a geographical dimension
3. Identify the total number of customers in each region in the United Kingdom
4. Create a pie graph identifying customer segmentation by region, gender
5. Represent the values in the pie chart as percentages using quick table calculation

Session 6 - Tableau for Data Science

6. What all quick table calculations you can do using Tableau?

Information reference:

onlinehelp.tableau.com/current/pro/desktop/en-us/calculations_tablecalculations_quick.html

7. How are customers distributed by age? Which visualization would you use to plot age distributions?

8. Create bins of size “5 years” to plot the age distributions using Histogram. Represent age in percentage terms on the Y- axis

9. Create an identical distribution for bank balance held by customers. Set bin size to 10,000 pounds and represent bank balance using percentage terms on the Y-axis

10. Parameters are used to control visualizations, create a parameter for age groups with 1 (minimum) and 10 (maximum) bin sizes and apply this parameter to age bin in the dimensions pane.

Information reference:

interworks.com/blog/anonymous/2012/03/26/how-create-and-use-parameters-tableau

11. Create a parameter for bank balance distribution with 1000 (minimum) and 25000 (maximum) bin sizes. Apply this parameter to bank balance bin in the dimension pane.

12. Create a treemap with job classification variable

13. At this point you would derive insights about the customer segments using dashboards. Create a dashboard with all the worksheets you have prepared

14. How do you think that each region differs in terms of bank balances and employment?

15. Which regions have highest representations by female account holders?

16. Which region has a younger demographic and predominately female account holders?

Session 6 - Tableau for Data Science

17. Do you observe any skew in the distributions from your customer data? If yes, what kind of insights do they provide?
18. What would be your recommendation for developing new products? let's say if it plans to sell life insurance, health insurance, and accident insurance products?
19. Which segments and regions should have targeted insurance services, say considering by age or by gender or employment?
20. Create a storyline compiling your analysis of the customers by bank balance, employment classification, age and other factors.

Session 6 - Tableau for Data Science

Case study nine

Objective: Tableau and R Integration

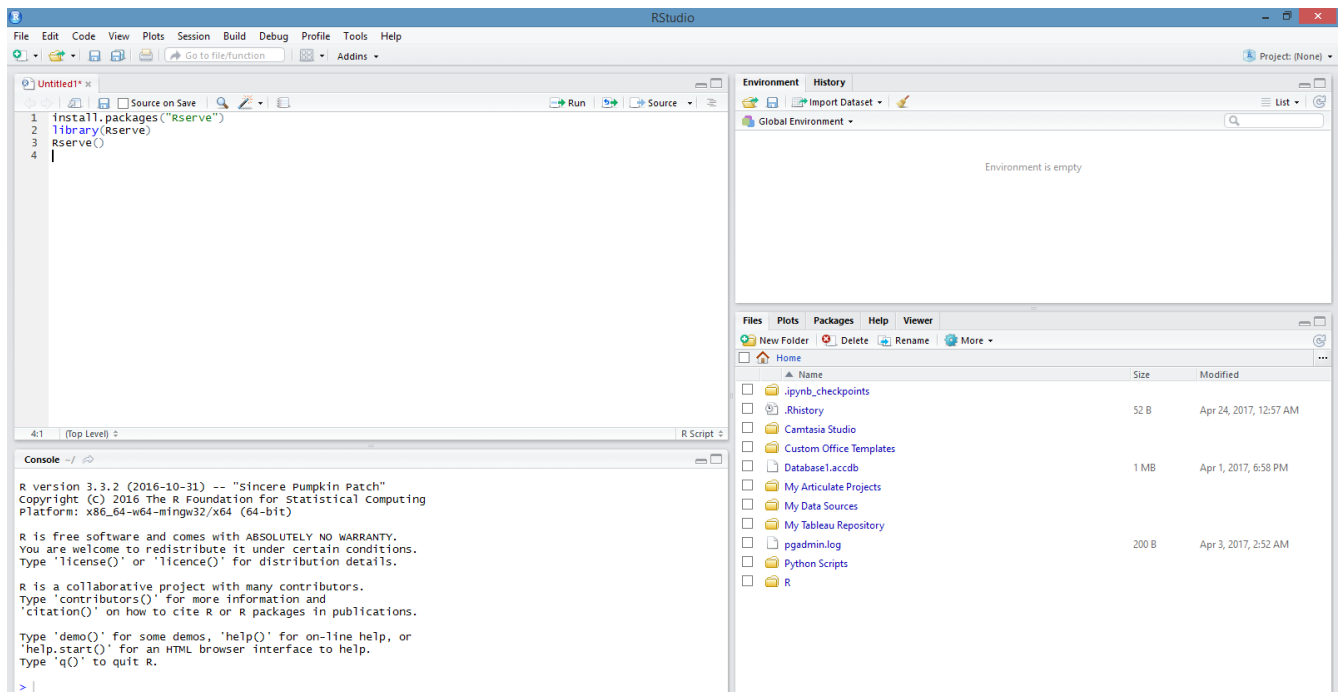
Step 1: First Install R Programming application and then install R - Studio desktop opensource edition

Download links

<https://cran.r-project.org/>

<https://www.rstudio.com/>

Step 2: Open R Studio, the Look of the R studio interface is displayed below



Session 6 - Tableau for Data Science

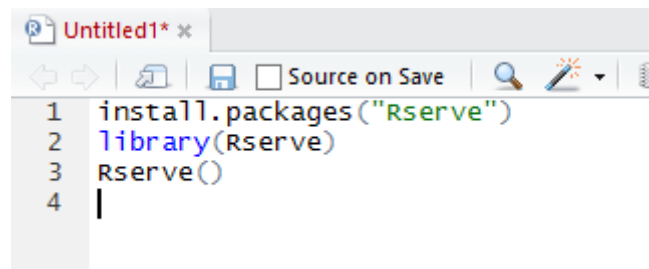
Step 3: install and start **Rserve** package

Use the following commands (execute them one at a time) by hitting run (ctrl + shift + enter)

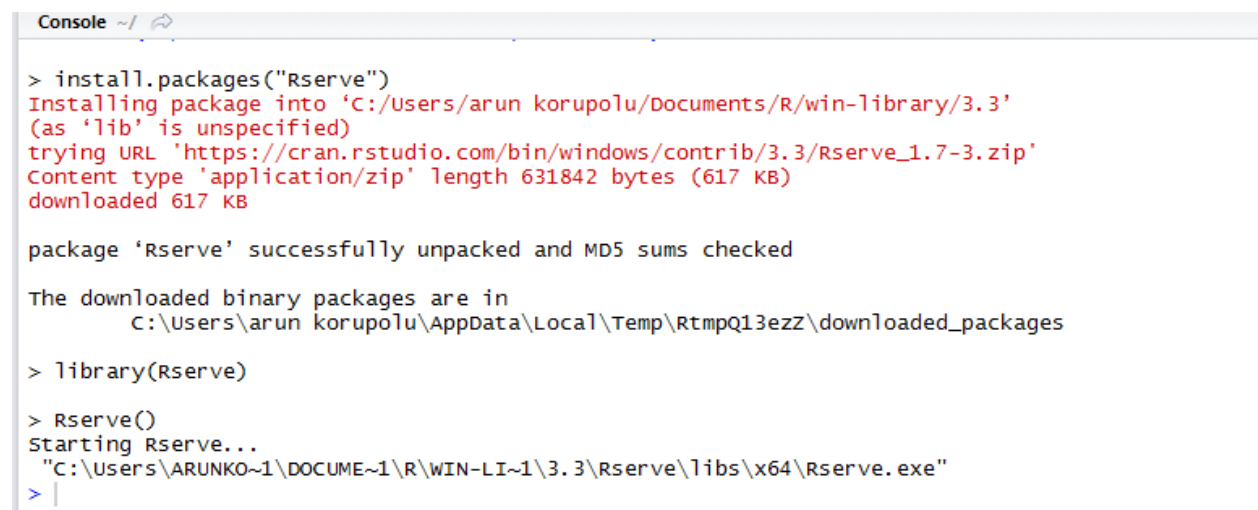
```
install.packages("Rserve")
```

```
library(Rserve)
```

```
Rserve()
```

A screenshot of the RStudio script editor window. The title bar says 'Untitled1* x'. The toolbar includes icons for back, forward, save, source on save, search, and a run button. The script contains four lines of code:

```
1 install.packages("Rserve")
2 library(Rserve)
3 Rserve()
4 |
```

A screenshot of the RStudio console window. It shows the output of the commands entered in the script editor. The output is as follows:

```
> install.packages("Rserve")
Installing package into 'C:/Users/arun korupolu/Documents/R/win-library/3.3'
(as 'lib' is unspecified)
trying URL 'https://cran.rstudio.com/bin/windows/contrib/3.3/Rserve_1.7-3.zip'
Content type 'application/zip' length 631842 bytes (617 KB)
downloaded 617 KB

package 'Rserve' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
      C:\Users\arun korupolu\AppData\Local\Temp\RtmpQ13ezz\downloaded_packages

> library(Rserve)

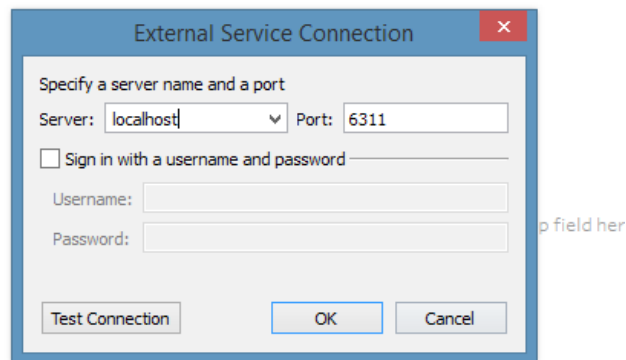
> Rserve()
Starting Rserve...
"C:\Users\ARUNKO~1\DOCUME~1\R\WIN-LI~1\3.3\Rserve\libs\x64\Rserve.exe"
> |
```

At this point minimize R studio window and start Tableau Desktop enterprise version

Step 4: open tableau desktop enterprise edition and check the connectivity

Goto: Help -> Settings & Performance -> Manage external service connection ☐ Click test connection

Session 6 - Tableau for Data Science



Step 5: Test whether the connection has been established

