

Using the Stargazer package in R

[The Stargazer package](#) for R provides a way to create publication quality tables, and a way for researchers to avoid creating new tables each time they tweak their dataset. This package saves users time, and has been [welcomed by the R community](#). It outputs tables in multiple formats; from .txt to LaTeX code as well as .html. This tutorial will go through the .txt and .html formats and provide the basic understanding needed to create Summary Statistics Tables and Regression Tables.

The RScript seen in this tutorial is available for download on the ECONPress website.

1. As with any R package, the first step is to install and load it. This can be done by typing “install.packages(“stargazer”)”, and then “library(stargazer)” in the next line. Installing Stargazer will only need to be done once, but the second command, which loads the package will need to be typed each session you wish to use it.

```
8
9 # Install the package, this is only required once
10 install.packages("stargazer")
11 # Load the package, this is required each use
12 library(stargazer)
13
```

## Summary Statistics

Summary Statistics tables typically provide sample population counts as well as averages for variables in the dataset. Stargazer’s default will produce a table with both of these measures as well as Standard Deviation, Minimum and Maximum values.

2. The syntax for the Summary Statistics command has four main arguments. The first specifies your dataframe (the example is using one called bull), the second says what type of output you’d like, the third a title for your table and the fifth finally a filename for your exported table. One example of this could look like “stargazer(bull, type = “text”, title=“Table 1: Summary Statistics”, out=“table1.txt”)”.

```
14 ### Summary Statistics
15 # Stargazer has four main Arugments
16 stargazer(bull, type = "text", title="Table 1: Summary statistics", out="table1.txt")
17
```

- a. The exported file is created in your working directory. While this example creates a .txt file that is visible in the Console window, replacing type with “html” and specifying “table1.html” in the out argument would create a HTML file that can be easily copied/pasted into Word/Excel. HTML output may look like meaningless code in the console.

```

Console ~/
> ### Summary Statistics
> # stargazer has four main Arugments
> stargazer(bull, type = "text", title="Table 1: Summary statistics", out="table1.txt")

Table 1: Summary statistics
=====
Statistic      N      Mean      St. Dev.      Min      Max
-----
Rank15         58    51.310     62.036         1      279
YearBorn       58   1,987.172     3.987       1,979   1,994
Height         58    68.638     2.367         60      76
Weight         58   153.828    14.033        115     188
YearsPro       58     7.948     3.753          1      17
Events14       58    15.328    10.300          1      28
BuckOuts14    58    43.034    31.044          1      93
Rides14        58    15.448    13.351          0      50
CupPoints14   58   3,631.448   2,452.617   136.750   9,520.250
Rank14         58    24.345    18.348          0      61
RidePer14     58     0.302     0.154         0.000   0.667
RidesPer_45bull_14 58    0.040     0.105         0.000   0.500
Rides90pts_14 58     0.552     1.029          0        5
Wins14         58     0.466     0.821          0        3
Top5_14        58     2.121     2.443          0       10
Top10_14       58     4.172     4.005          0       15
FinalPoints14 58    503.431    1,027.466     0.000   4,553.500
Earnings14     58  126,468.100  197,309.500  1,000.000  1,422,603.000
Events13       58    12.655    11.064          0       27
BuckOuts13    58    36.052    32.584          0       91
Rides13        58    13.241    13.701          0       50
CupPoints13   58   3,109.197   2,959.051     0.000  10,937.750
Rank13         58    13.966    15.269          0       50
RidePer13     58     0.229     0.197         0.000   0.579
RidesPer_45bull_13 58    0.051     0.161         0.000   1.000
Rides90pts_13 58     0.517     1.274          0        8
Wins13         58     0.448     0.976          0        5
Top5_13        58     1.931     2.771          0       11
Top10_13       58     3.810     4.190          0       14
FinalPoints13 58    508.556    1,067.353     0.000   5,296.250
Earnings13     58  115,248.500  252,794.400     0.000  1,810,711.000
Events12       58    10.397    11.974          0       29
BuckOuts12    58    30.466    36.128          0      103
Rides12        58    13.345    17.867          0       62
CupPoints12   58   2,764.095   3,856.021     0.000  12,201.750
Rank12         58    10.569    15.940          0       65
RidePer12     58     0.199     0.220         0.000   0.610
Wins12         58     0.414     0.726          0        3
Top5_12        58     1.983     3.092          0       10
Top10_12       58     3.655     4.930          0       18
FinalPoints12 58    384.634    913.585     0.000   4,189.250
Earnings12     58  100,685.400  213,028.900     0.000  1,464,476.000
=====

```

- To only display a subset of your dataframe, add a vector to your first argument specifying the variables you'd like to include.

```

20
21 # This notation provides summary statistics for just the provided variables.
22 stargazer(bull[c("Rank14","Rank13","Rank12")], type = "text",
23           title="Table 1: Summary statistics for selected variables", out="table2.txt")
24

```

- What used to just be "dataname," would become "dataname[c("var1","var2","etc")),"

```
> # This notation provides summary statistics for just the provided variables.
> stargazer(bull[c("Rank14","Rank13","Rank12")], type = "text",
+ title="Table 1: Summary statistics for selected variables", out="table2.txt")
```

Table 1: Summary statistics for selected variables

Statistic	N	Mean	St. Dev.	Min	Max
Rank14	58	24.345	18.348	0	61
Rank13	58	13.966	15.269	0	50
Rank12	58	10.569	15.940	0	65

4. Adding “,flip=TRUE” inverts your column/rows so that variables are the column headers.

```
24
25 # Adding the "flip=TRUE" argument transposes the information so variables are in columns.
26 stargazer(bull[c("Rank14","Rank13","Rank12")], type = "text",
27           title="Table 1: Summary statistics", out="table3.txt", flip=TRUE)
28
```

```
> # Adding the "flip=TRUE" argument transposes the information so variables are in columns.
> stargazer(bull[c("Rank14","Rank13","Rank12")], type = "text",
+ title="Table 1: Summary statistics", out="table3.txt", flip=TRUE)
```

Table 1: Summary statistics

Statistic	Rank14	Rank13	Rank12
N	58	58	58
Mean	24.345	13.966	10.569
St. Dev.	18.348	15.269	15.940
Min	0	0	0
Max	61	50	65

5. Adding “,digits=1” controls the number of decimal places displayed to only show one. The default is to show 3 decimal places.

```
28
29 # Adding the "digits" argument allows the user to control the number of decimals displayed.
30 stargazer(bull[c("Rank14","Rank13","Rank12")], type = "text",
31           title="Table 1: Summary statistics", out="table4.txt", flip=TRUE, digits=1)
32
```

```
> # Adding the "digits" argument allows the user to control the number of decimals displayed.
> stargazer(bull[c("Rank14","Rank13","Rank12")], type = "text",
+ title="Table 1: Summary statistics", out="table4.txt", flip=TRUE, digits=1)
```

Table 1: Summary statistics

Statistic	Rank14	Rank13	Rank12
N	58	58	58
Mean	24.3	14.0	10.6
St. Dev.	18.3	15.3	15.9
Min	0	0	0
Max	61	50	65

6. To replace the variable names from your dataframe with custom variable names, simply create a vector of variable of names and add it with the “covariate.labels” argument. This vector will

relabel your variables in the exact order provided, so be sure your vector corresponds to the order of your variables.

```

32
33 # Adding the "covariate.labels" argument gives custom names to your variables.
34 # Note that it doesn't match at all, just relabels them in the order of the provided vector.
35 stargazer(bull[c("Rank14", "Rank13", "Rank12")], type = "text",
36           title="Table 1: Summary statistics", out="table5.txt", digits=1,
37           covariate.labels=c("2014 Rider Rank", "2013 Rider Rank", "2012 Rider Rank"))
38

```

```

> # Adding the "covariate.labels" argument gives custom names to your variables.
> # Note that it doesn't match, just relabels them in the order of the provided vector.
> stargazer(bull[c("Rank14", "Rank13", "Rank12")], type = "text",
+ title="Table 1: Summary statistics", out="table5.txt", digits=1,
+ covariate.labels=c("2014 Rider Rank", "2013 Rider Rank", "2012 Rider Rank"))

```

Table 1: Summary statistics

Statistic	N	Mean	St. Dev.	Min	Max
2014 Rider Rank	58	24.3	18.3	0	61
2013 Rider Rank	58	14.0	15.3	0	50
2012 Rider Rank	58	10.6	15.9	0	65

7. Stargazer allows you to show summary statistics for a subset of your data, preventing you from needing to create and manage multiple dataframes. To do this, preface your dataframe with "subset(" and add a condition before you close the parentheses. In this example, we're commanding Stargazer to only show us summary statistics for riders over 67 inches tall.

```

38
39 # Summary statistics for a subset of the data.
40 stargazer(subset(bull[c("Rank14", "Rank13", "Rank12")], bull$Height>=67),
41           title="Select Summary Statistics for Riders over 65inches",
42           type = "text", digits=1, out="table6.txt")
43

```

```

> # Summary statistics for a subset of the data.
> stargazer(subset(bull[c("Rank14", "Rank13", "Rank12")], bull$Height>=67),
+ title="select summary statistics for Riders over 65inches",
+ type = "text", digits=1, out="table6.txt")

```

Select Summary Statistics for Riders over 65inches

Statistic	N	Mean	St. Dev.	Min	Max
Rank14	49	25.1	18.9	0	61
Rank13	49	14.0	15.5	0	50
Rank12	49	10.7	16.8	0	65

## Regression Output

Developing regression tables with Stargazer just requires that you name the results and provide them alongside the aforementioned arguments.

1. Set up, run and name your regressions as would be normal for R (here we have made a subset of our data just to ensure only relevant observations are being used).
2. Then list your regression names and again specify the type of table you would like to export, list custom variable names, choose how many decimals you'd like to be displayed etc.
3. You can see that Stargazer creates a publication quality table for you to make use of. The HTML option creates a file that is easily copy/pasted into word or excel for further formatting.
4. Stargazer provides many more options for creating your regression tables so be sure to check out the [official R-Project write up for Stargazer](#) to learn more.

```

44
45 ### Regression Tables
46 bull_14 <- bull[bull$Events14 >0,]
47 # Run and name regressions as normal
48 m1 <- lm(Earnings14 ~ RidePer14, data=bull_14)
49 m2 <- lm(Earnings14 ~ RidePer14 + Events14, data=bull_14)
50 m3 <- lm(Earnings14 ~ RidePer14 + Events14 + BuckOuts14, data=bull_14)
51 |
52 # Follow the same general format as above, with regression names instead of a dataframe
53 # use the "dep.var.labels" argument to label your various dependent variables
54 stargazer(m1, m2, m3, type="text",dep.var.labels=c("2014 Earnings"),
55           title="Table 2: Regression Results",digits=1,out="models.txt",
56           covariate.labels=c("Riding Percentage","# Events","Buck Outs"))
57

```

```

> ### Regression Tables
> bull_14 <- bull[bull$Events14 >0,]
> # Run and name regressions as normal
> m1 <- lm(Earnings14 ~ RidePer14, data=bull_14)
> m2 <- lm(Earnings14 ~ RidePer14 + Events14, data=bull_14)
> m3 <- lm(Earnings14 ~ RidePer14 + Events14 + BuckOuts14, data=bull_14)
> # Follow the same general format as above, with regression names instead of a dataframe
> # Use the "dep.var.labels" argument to label your various dependent variables
> stargazer(m1, m2, m3, type="text",dep.var.labels=c("2014 Earnings"),
+ title="Table 2: Regression Results",digits=1,out="models.txt",
+ covariate.labels=c("Riding Percentage","# Events","Buck Outs"))

```

Table 2: Regression Results

	Dependent variable:		
	(1)	2014 Earnings (2)	(3)
Riding Percentage	485,119.2*** (157,951.2)	270,725.3 (172,419.9)	73,811.3 (161,727.7)
# Events		6,612.0** (2,585.2)	-50,672.8*** (14,850.9)
Buck Outs			19,674.2*** (5,038.7)
Constant	-20,088.1 (53,494.7)	-56,664.5 (52,995.4)	34,191.1 (52,651.3)
Observations	58	58	58
R2	0.1	0.2	0.4
Adjusted R2	0.1	0.2	0.4
Residual Std. Error	184,156.5 (df = 56)	175,669.6 (df = 55)	156,560.0 (df = 54)
F Statistic	9.4*** (df = 1; 56)	8.5*** (df = 2; 55)	12.2*** (df = 3; 54)

Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01