Creating Data Frames using R

Randy Richeson

Objectives

- What is R?
- Basic Data Structures in R
- Getting Started
- Creating Vectors
- Creating Data Frames
- Using Data Frames
- Plotting Data Frames

What is R?

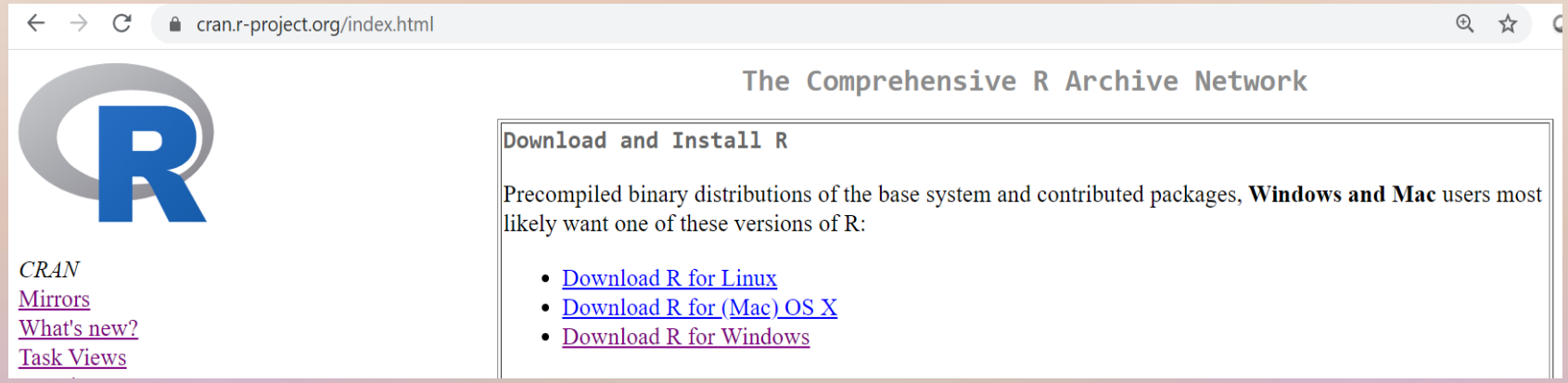
- An open-source language and environment for statistical analysis and graphics.
- Available under the Free Software Foundation's GNU General Public License in source code form.
- Provides more than 10,000 packages with functions including statistical algorithms used for linear and nonlinear modeling, classical, time-series, classification, and clustering analyses
- Provides many graphing techniques 
- Compiles and runs on a wide variety of operating systems

Basic Data Structures in R

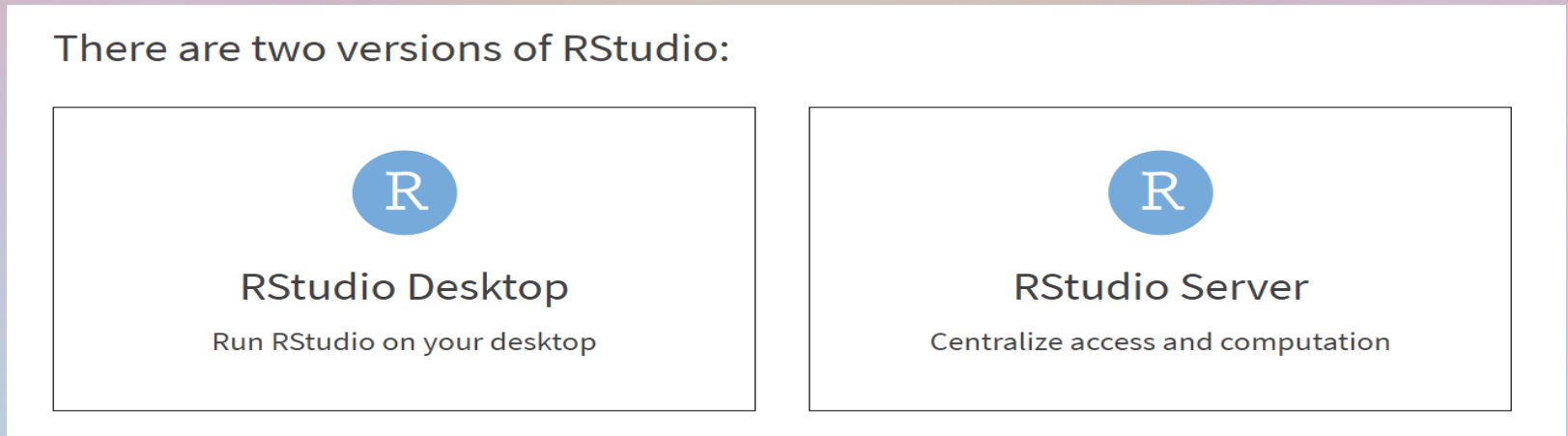
- **Vector**: Ordered **c**ollection of character, logical, integer or numeric values 
- List: Ordered collection of objects with elements that encompass any mixture of data types and can be named.
- Matrix**:** Vector with dimensions
- **Data Frame**: A special type of list where every element of the list has the same length 
- Factor. Vector of objects with a discrete grouping of components
- Table. Similar to an array

Getting Started

- Download R from CRAN



- Optionally, download your favorite IDE



Creating Vectors

c Combines Values Into A Vector Or List

```
> id = c (1,2,3,4,5)
> name = c("Mt. Elbert","Mt. Massive","Mt. Harvard","Blanca Peak","La Plata Peak")
> feet = c(14433,14421,14420,14345,14336)
> ascents = c(12611,8915,6859,4525,8485)
```

OR

Vectors can be used to create data frames

```
> id <- c (1,2,3,4,5)
> name <- c("Mt. Elbert","Mt. Massive","Mt. Harvard","Blanca Peak","La Plata Peak")
> feet <- c(14433,14421,14420,14345,14336)
> ascents <- c(12611,8915,6859,4525,8485)
```

Creating Data Frames Using Vectors

```
>  CO14ers = data.frame(id=id,name=name,feet=feet,ascents=ascents)
```

OR

```
>  CO14ers <- data.frame(id=id,name=name,feet=feet,ascents=ascents)
```

OR

```
> data.frame(id=id,name=name,feet=feet,ascents=ascents) -> CO14ers
```

Creating Data Frames Using CSV File

```
> fourteeners <- read.csv("C:/job_search/appian/fourteeners.csv")
```

OR

```
> read.csv("C:/job_search/appian/fourteeners.csv") -> fourteeners
```

read.csv reads a file and returns a data frame

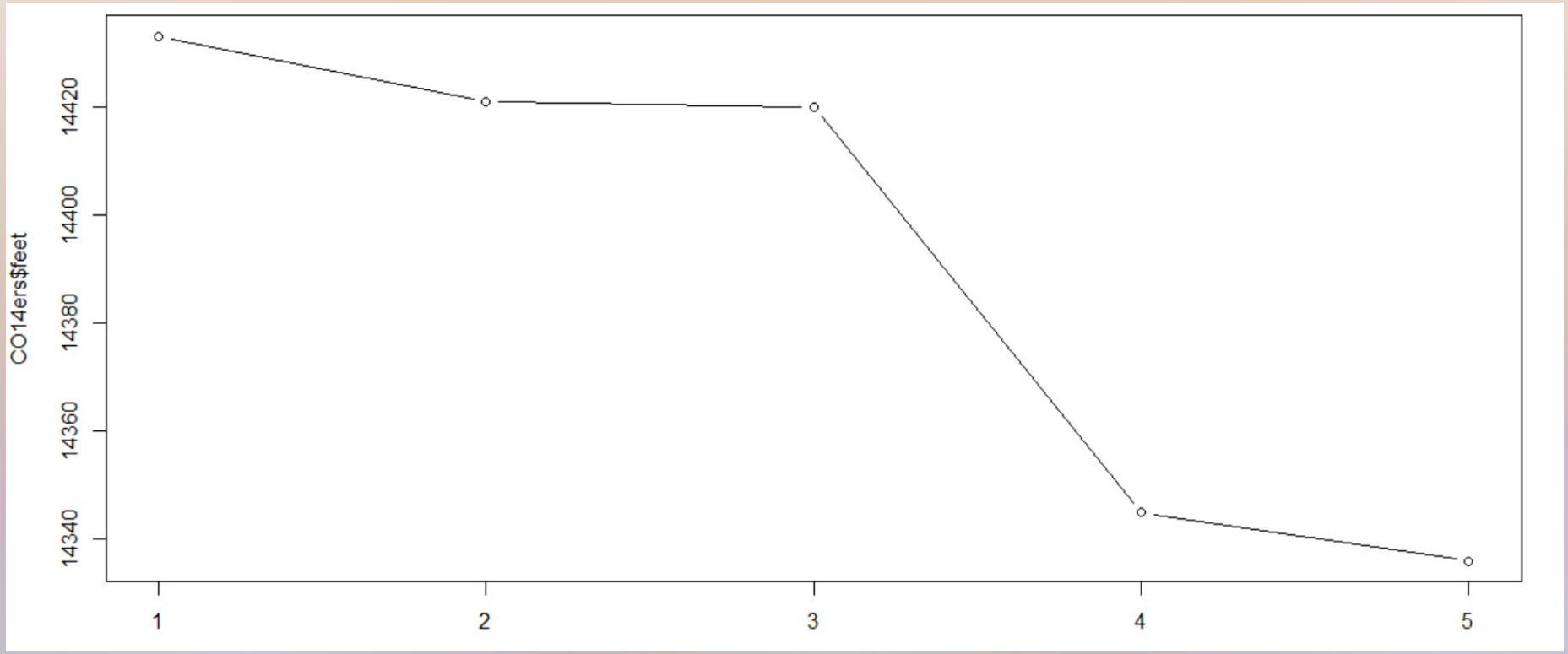
Using a Data Frame

```
> CO14ers
   id  name            feet        ascents
1  1    Mt. Elbert      14433       12611
2  2    Mt. Massive     14421       8915
3  3    Mt. Harvard     14420       6859
4  4    Blanca Peak     14345       4525
5  5    La Plata Peak   14336       8485

> mean(CO14ers$feet)
[1] 14391
```

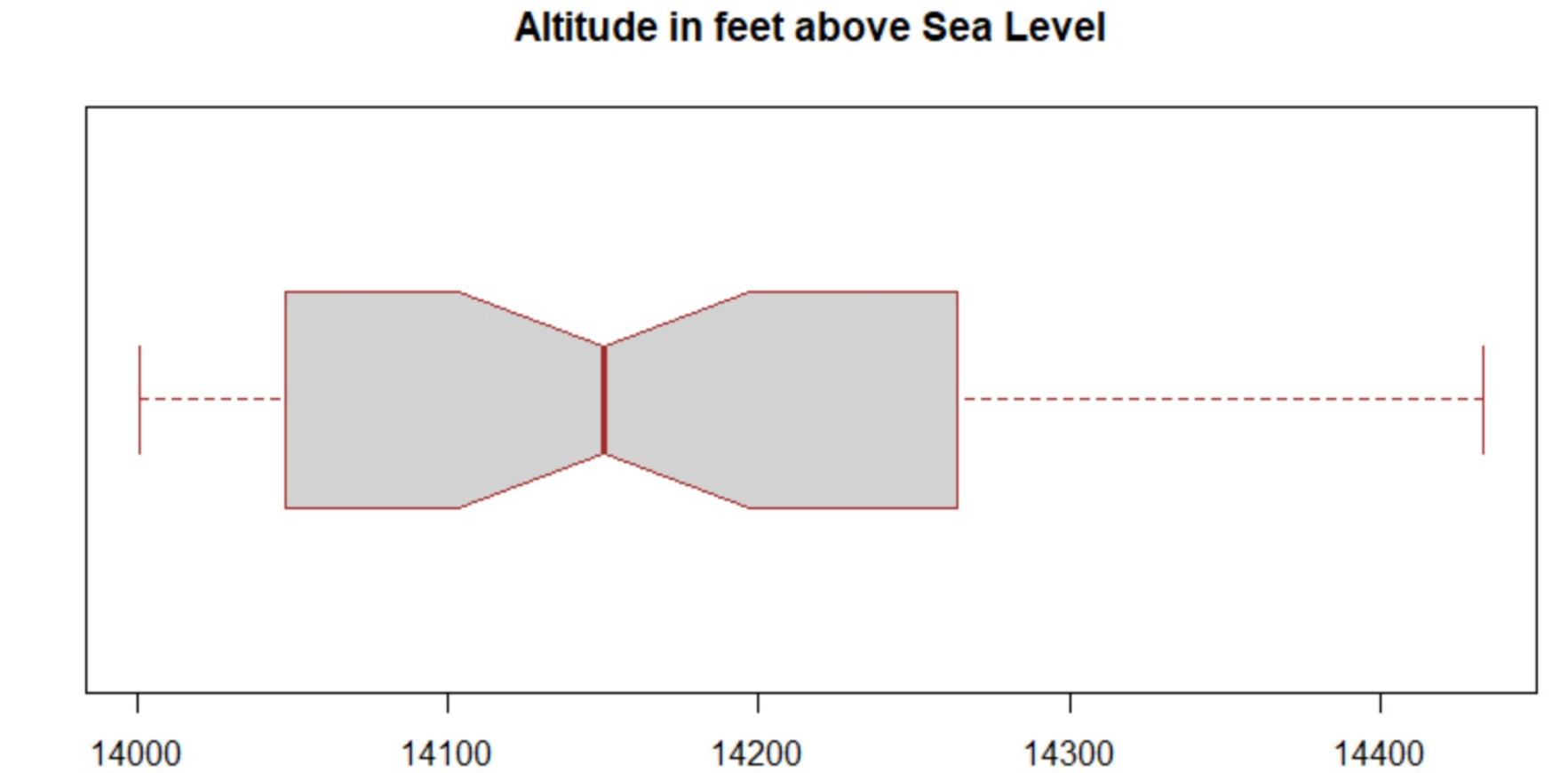
Plotting a Data Frame

```
> plot(CO14ers$feet, type="b")
```



Using a Box Plot with Data Frame

> boxplot(fourteeners$Elevation, main = "Altitude in feet above Sea Level", xlab = "Elevation", border = "brown", horizontal = TRUE, notch = TRUE)



Questions ???