

Reading In Datasets

Ben Goldstone

Jan 25, 2023

Reading in Textbook Data

```
data("StudentSurvey")
head(StudentSurvey)
```

	Year	Sex	Smoke	Award	HigherSAT	Exercise	TV	Height	Weight	Siblings
## 1	Senior	M	No	Olympic	Math	10	1	71	180	4
## 2	Sophomore	F	Yes	Academy	Math	4	7	66	120	2
## 3	FirstYear	M	No	Nobel	Math	14	5	72	208	2
## 4	Junior	M	No	Nobel	Math	3	1	63	110	1
## 5	Sophomore	F	No	Nobel	Verbal	3	3	65	150	1
## 6	Sophomore	F	No	Nobel	Verbal	5	4	65	114	2

	BirthOrder	VerbalSAT	MathSAT	SAT	GPA	Pulse	Piercings
## 1	4	540	670	1210	3.13	54	0
## 2	2	520	630	1150	2.50	66	3
## 3	1	550	560	1110	2.55	130	0
## 4	1	490	630	1120	3.10	78	0
## 5	1	720	450	1170	2.70	40	6
## 6	2	600	550	1150	3.20	80	4

It looks like this dataset is demographic data on students based on academic performance, family, health, and awards.

Reading in a csv file dataset

```
library(readr)
Quarterbacks <- read_csv("CSVs/Quarterbacks.csv")
```

```
## Rows: 251 Columns: 20
## -- Column specification -----
## Delimiter: ","
## chr (2): Player, HOF
## dbl (18): HOFScore, ChampionshipWins, AllProSelections, ProBowlSelections, Y...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
head(Quarterbacks)
```

```
## # A tibble: 6 x 20
##   Player          HOFScore ChampionshipWins AllProSelections ProBowlSelections
```

```
##   <chr>           <dbl>           <dbl>           <dbl>           <dbl>
## 1 Tom Brady      259.             7             3             15
## 2 Peyton Manning 258.             2             7             14
## 3 Brett Favre    180.             1             3             11
## 4 Aaron Rodgers  175.             1             4             10
## 5 Johnny Unitas  160.             3             5             10
## 6 Joe Montana    154.             4             3             8
## # ... with 15 more variables: YearsStarter <dbl>, FirstYearStarter <dbl>,
## #   LastYearStarter <dbl>, WeightedCareerAve <dbl>, GamesPlayed <dbl>,
## #   GamesStarted <dbl>, PassesCompleted <dbl>, PassesAttempted <dbl>,
## #   YardsGained <dbl>, PassingTouchdowns <dbl>, LongestPass <dbl>,
## #   Interceptions <dbl>, Sacked <dbl>, YardsLostSacks <dbl>, HOF <chr>
```

This dataset looks like it is collecting data on on NFL Quarterbacks