

## R Notebook : Assignment 2

```
library(ggplot2)
library(readxl)

collegedata.df <- read_excel(file.choose())
head(collegedata.df)

## # A tibble: 6 x 130
##   Profile                                Name year `ZIP code`
##   <chr>                                <chr> <dbl>    <chr>
## 1 Other      York College Pennsylvania  2013 17403-3651
## 2 Other      Yeshiva University          2013 10033-3299
## 3 Other      Yale University                2013      6520
## 4 Other      Xavier University of Louisiana  2013 70125-1098
## 5 Other      Xavier University                2013 45207-1092
## 6 Other      Winston-Salem State University  2013 27110-0001
## # ... with 126 more variables: `Highest degree offered` <chr>, `County
## #   name` <chr>, `Longitude location of institution` <dbl>, `Latitude
## #   location of institution` <dbl>, `Religious affiliation` <chr>, `Offers
## #   Less than one year certificate` <chr>, `Offers One but less than two
## #   years certificate` <chr>, `Offers Associate's degree` <chr>, `Offers
## #   Two but less than 4 years certificate` <chr>, `Offers Bachelor's
## #   degree` <chr>, `Offers Postbaccalaureate certificate` <chr>, `Offers
## #   Master's degree` <chr>, `Offers Post-master's certificate` <chr>,
## #   `Offers Doctor's degree - research/scholarship` <chr>, `Offers
## #   Doctor's degree - professional practice` <chr>, `Offers Doctor's
## #   degree - other` <chr>, `Offers Other degree` <chr>, `Applicants
## #   total` <dbl>, `Admissions total` <dbl>, `Enrolled total` <dbl>,
## #   `Percent of freshmen submitting SAT scores` <dbl>, `Percent of
## #   freshmen submitting ACT scores` <dbl>, `ACT Composite 25th percentile
## #   score` <dbl>, `ACT Composite 75th percentile score` <dbl>, `Estimated
## #   enrollment, total` <dbl>, `Estimated enrollment, full time` <dbl>,
## #   `Estimated enrollment, part time` <dbl>, `Estimated undergraduate
## #   enrollment, total` <dbl>, `Estimated undergraduate enrollment, full
## #   time` <dbl>, `Estimated undergraduate enrollment, part time` <dbl>,
## #   `Estimated freshman undergraduate enrollment, total` <dbl>, `Estimated
## #   freshman enrollment, full time` <dbl>, `Estimated freshman enrollment,
## #   part time` <dbl>, `Estimated graduate enrollment, total` <dbl>,
## #   `Estimated graduate enrollment, full time` <dbl>, `Estimated graduate
## #   enrollment, part time` <dbl>, `Associate's degrees awarded` <dbl>,
## #   `Bachelor's degrees awarded` <dbl>, `Master's degrees awarded` <dbl>,
## #   `Doctor's degree - research/scholarship awarded` <dbl>, `Doctor's
## #   degrees - professional practice awarded` <dbl>, `Doctor's degrees -
## #   other awarded` <dbl>, `Certificates of less than 1-year
## #   awarded` <dbl>, `Certificates of 1 but less than 2-years
## #   awarded` <dbl>, `Certificates of 2 but less than 4-years
```

```

## # awarded` <dbl>, `Postbaccalaureate certificates awarded` <dbl>,
## # `Post-master's certificates awarded` <dbl>, `Number of students
## # receiving an Associate's degree` <dbl>, `Number of students receiving
## # a Bachelor's degree` <dbl>, `Number of students receiving a Master's
## # degree` <dbl>, `Number of students receiving a Doctor's degree` <dbl>,
## # `Number of students receiving a certificate of less than
## # 1-year` <dbl>, `Number of students receiving a certificate of 1 but
## # less than 4-years` <dbl>, `Number of students receiving a
## # Postbaccalaureate or Post-master's certificate` <dbl>, `Percent
## # admitted - total` <dbl>, `Admissions yield - total` <dbl>, `Tuition
## # and fees, 2010-11` <dbl>, `Tuition and fees, 2011-12` <dbl>, `Tuition
## # and fees, 2012-13` <dbl>, `Tuition and fees, 2013-14` <dbl>, `Total
## # price for in-state students living on campus 2013-14` <dbl>, `Total
## # price for out-of-state students living on campus 2013-14` <dbl>,
## # `State abbreviation` <chr>, `FIPS state code` <chr>, `Geographic
## # region` <chr>, `Sector of institution` <chr>, `Level of
## # institution` <chr>, `Control of institution` <chr>, `Historically
## # Black College or University` <chr>, `Tribal college` <chr>, `Degree of
## # urbanization (Urban-centric locale)` <chr>, `Carnegie Classification
## # 2010: Basic` <chr>, `Total enrollment` <dbl>, `Full-time
## # enrollment` <dbl>, `Part-time enrollment` <dbl>, `Undergraduate
## # enrollment` <dbl>, `Graduate enrollment` <dbl>, `Full-time
## # undergraduate enrollment` <dbl>, `Part-time undergraduate
## # enrollment` <dbl>, `Percent of total enrollment that are American
## # Indian or Alaska Native` <dbl>, `Percent of total enrollment that are
## # Asian` <dbl>, `Percent of total enrollment that are Black or African
## # American` <dbl>, `Percent of total enrollment that are
## # Hispanic/Latino` <dbl>, `Percent of total enrollment that are Native
## # Hawaiian or Other Pacific Islander` <dbl>, `Percent of total
## # enrollment that are White` <dbl>, `Percent of total enrollment that
## # are two or more races` <dbl>, `Percent of total enrollment that are
## # Race/ethnicity unknown` <dbl>, `Percent of total enrollment that are
## # Nonresident Alien` <dbl>, `Percent of total enrollment that are
## # Asian/Native Hawaiian/Pacific Islander` <dbl>, `Percent of total
## # enrollment that are women` <dbl>, `Percent of undergraduate enrollment
## # that are American Indian or Alaska Native` <dbl>, `Percent of
## # undergraduate enrollment that are Asian` <dbl>, `Percent of
## # undergraduate enrollment that are Black or African American` <dbl>,
## # `Percent of undergraduate enrollment that are Hispanic/Latino` <dbl>,
## # `Percent of undergraduate enrollment that are Native Hawaiian or Other
## # Pacific Islander` <dbl>, `Percent of undergraduate enrollment that are
## # White` <dbl>, `Percent of undergraduate enrollment that are two or
## # more races` <dbl>, `Percent of undergraduate enrollment that are
## # Race/ethnicity unknown` <dbl>, `Percent of undergraduate enrollment
## # that are Nonresident Alien` <dbl>, `Percent of undergraduate
## # enrollment that are Asian/Native Hawaiian/Pacific Islander` <dbl>, ...

```

```
str(collegedata.df)
```

```

## Classes 'tbl_df', 'tbl' and 'data.frame':    570 obs. of  130 variables:
## $ Profile
: chr  "Other" "Other" "Other" "Other" ...
## $ Name
: chr  "York College Pennsylvania" "Yeshiva University" "Yale University"
"Xavier University of Louisiana" ...
## $ year
: num  2013 2013 2013 2013 2013 ...
## $ ZIP code
: chr  "17403-3651" "10033-3299" "6520" "70125-1098" ...
## $ Highest degree offered
: chr  "Doctor's degree - professional practice" "Doctor's degree -
research/scholarship and professional practice" "Doctor's degree -
research/scholarship and professional practice" "Doctor's degree -
professional practice" ...
## $ County name
: chr  "York County" "New York County" "New Haven County" "Orleans Parish"
...
## $ Longitude location of institution
: num  -76.7 -73.9 -72.9 -90.1 -84.5 ...
## $ Latitude location of institution
: num  39.9 40.9 41.3 30 39.1 ...
## $ Religious affiliation
: chr  "Not applicable" "Not applicable" "Not applicable" "Roman Catholic"
...
## $ Offers Less than one year certificate
: chr  "Implied no" "Implied no" "Implied no" "Implied no" ...
## $ Offers One but less than two years certificate
: chr  "Implied no" "Implied no" "Implied no" "Implied no" ...
## $ Offers Associate's degree
: chr  "Yes" "Yes" "Implied no" "Implied no" ...
## $ Offers Two but less than 4 years certificate
: chr  "Implied no" "Implied no" "Implied no" "Implied no" ...
## $ Offers Bachelor's degree
: chr  "Yes" "Yes" "Yes" "Yes" ...
## $ Offers Postbaccalaureate certificate
: chr  "Implied no" "Implied no" "Yes" "Implied no" ...
## $ Offers Master's degree
: chr  "Yes" "Yes" "Yes" "Yes" ...
## $ Offers Post-master's certificate
: chr  "Yes" "Yes" "Yes" "Implied no" ...
## $ Offers Doctor's degree - research/scholarship
: chr  "Implied no" "Yes" "Yes" "Implied no" ...
## $ Offers Doctor's degree - professional practice
: chr  "Yes" "Yes" "Yes" "Yes" ...
## $ Offers Doctor's degree - other
: chr  "Implied no" "Implied no" "Yes" "Implied no" ...
## $ Offers Other degree
: chr  "Implied no" "Implied no" "Implied no" "Implied no" ...
## $ Applicants total

```

```

: num  9934 1575 28977 4703 10907 ...
## $ Admissions total
: num  7302 1294 2043 2560 7632 ...
## $ Enrolled total
: num  1064 857 1356 675 1289 ...
## $ Percent of freshmen submitting SAT scores
: num  95 66 84 30 45 93 1 86 16 0 ...
## $ Percent of freshmen submitting ACT scores
: num  18 36 35 88 80 25 97 47 84 98 ...
## $ ACT Composite 25th percentile score
: num  20 23 32 20 23 16 21 19 19 20 ...
## $ ACT Composite 75th percentile score
: num  25 29 35 26 27 19 25 24 26 28 ...
## $ Estimated enrollment, total
: num  5248 6620 11960 3121 6671 ...
## $ Estimated enrollment, full time
: num  4492 5528 11800 2942 4811 ...
## $ Estimated enrollment, part time
: num  756 1092 160 179 1860 ...
## $ Estimated undergraduate enrollment, total
: num  5008 2850 5310 2351 4652 ...
## $ Estimated undergraduate enrollment, full time
: num  4453 2750 5300 2223 4150 ...
## $ Estimated undergraduate enrollment, part time
: num  555 100 10 128 502 619 897 56 107 372 ...
## $ Estimated freshman undergraduate enrollment, total
: num  1064 873 1300 675 1289 ...
## $ Estimated freshman enrollment, full time
: num  1063 871 1300 674 1284 ...
## $ Estimated freshman enrollment, part time
: num  1 2 NA 1 5 2 12 0 2 4 ...
## $ Estimated graduate enrollment, total
: num  240 3770 6650 770 2019 ...
## $ Estimated graduate enrollment, full time
: num  39 2778 6500 719 661 ...
## $ Estimated graduate enrollment, part time
: num  201 992 150 51 1358 ...
## $ Associate's degrees awarded
: num  26 291 0 0 18 0 37 0 10 0 ...
## $ Bachelor's degrees awarded
: num  975 585 1308 326 1033 ...
## $ Master's degrees awarded
: num  64 458 1618 44 803 ...
## $ Doctor's degree - research/scholarship awarded
: num  0 131 398 0 0 0 6 18 8 0 ...
## $ Doctor's degrees - professional practice awarded
: num  4 595 360 145 15 0 0 63 0 0 ...
## $ Doctor's degrees - other awarded
: num  0 0 0 0 0 0 0 0 0 0 ...
## $ Certificates of less than 1-year awarded

```

```

: num 0 0 0 0 0 0 0 0 0 0 ...
## $ Certificates of 1 but less than 2-years awarded
: num 0 0 0 0 0 0 0 0 0 0 ...
## $ Certificates of 2 but less than 4-years awarded
: num 0 0 0 0 0 0 0 0 0 0 ...
## $ Postbaccalaureate certificates awarded
: num 0 0 0 0 0 0 19 0 0 0 ...
## $ Post-master's certificates awarded
: num 2 11 500 0 0 1 3 30 143 0 ...
## $ Number of students receiving an Associate's degree
: num 26 291 0 0 18 0 37 0 10 0 ...
## $ Number of students receiving a Bachelor's degree
: num 975 585 1308 326 1015 ...
## $ Number of students receiving a Master's degree
: num 64 458 1594 44 800 ...
## $ Number of students receiving a Doctor's degree
: num 4 726 688 145 15 0 6 81 8 0 ...
## $ Number of students receiving a certificate of less than 1-year
: num 0 0 0 0 0 0 0 0 0 0 ...
## $ Number of students receiving a certificate of 1 but less than 4-years
: num 0 0 0 0 0 0 0 0 0 0 ...
## $ Number of students receiving a Postbaccalaureate or Post-master's
certificate : num 2 11 497 0 0 1 22 30 143 0 ...
## $ Percent admitted - total
: num 74 82 7 54 70 56 61 79 75 49 ...
## $ Admissions yield - total
: num 15 66 66 26 17 24 38 15 31 48 ...
## $ Tuition and fees, 2010-11
: num 15140 33050 38300 17100 29970 ...
## $ Tuition and fees, 2011-12
: num 15880 35200 40500 17900 31160 ...
## $ Tuition and fees, 2012-13
: num 16520 36500 42300 18700 32070 ...
## $ Tuition and fees, 2013-14
: num 17010 37600 44000 20560 33000 ...
## $ Total price for in-state students living on campus 2013-14
: num 29790 55075 61620 31501 46520 ...
## $ Total price for out-of-state students living on campus 2013-14
: num 29790 55075 61620 31501 46520 ...
## $ State abbreviation
: chr "Pennsylvania" "New York" "Connecticut" "Louisiana" ...
## $ FIPS state code
: chr "Pennsylvania" "New York" "Connecticut" "Louisiana" ...
## $ Geographic region
: chr "Mid East DE DC MD NJ NY PA" "Mid East DE DC MD NJ NY PA" "New England
CT ME MA NH RI VT" "Southeast AL AR FL GA KY LA MS NC SC TN VA WV" ...
## $ Sector of institution
: chr "Private not-for-profit, 4-year or above" "Private not-for-profit, 4-
year or above" "Private not-for-profit, 4-year or above" "Private not-for-
profit, 4-year or above" ...

```

```

## $ Level of institution
: chr "Four or more years" "Four or more years" "Four or more years" "Four
or more years" ...
## $ Control of institution
: chr "Private not-for-profit" "Private not-for-profit" "Private not-for-
profit" "Private not-for-profit" ...
## $ Historically Black College or University
: chr "No" "No" "No" "Yes" ...
## $ Tribal college
: chr "No" "No" "No" "No" ...
## $ Degree of urbanization (Urban-centric locale)
: chr "Suburb: Midsize" "City: Large" "City: Midsize" "City: Large" ...
## $ Carnegie Classification 2010: Basic
: chr "Master's Colleges and Universities (smaller programs)" "Research
Universities (very high research activity)" "Research Universities (very high
research activity)" "Baccalaureate Colleges--Arts & Sciences" ...
## $ Total enrollment
: num 5250 6747 12109 3121 6633 ...
## $ Full-time enrollment
: num 4492 5585 11927 2942 4811 ...
## $ Part-time enrollment
: num 758 1162 182 179 1822 ...
## $ Undergraduate enrollment
: num 5010 2887 5430 2504 4620 ...
## $ Graduate enrollment
: num 240 3860 6679 617 2013 ...
## $ Full-time undergraduate enrollment
: num 4453 2782 5424 2375 4150 ...
## $ Part-time undergraduate enrollment
: num 557 105 6 129 470 621 899 56 159 371 ...
## $ Percent of total enrollment that are American Indian or Alaska Native
: num 0 0 0 0 0 0 0 1 0 0 ...
## $ Percent of total enrollment that are Asian
: num 1 4 13 13 3 1 2 2 1 3 ...
## $ Percent of total enrollment that are Black or African American
: num 5 2 5 71 9 70 2 14 3 27 ...
## $ Percent of total enrollment that are Hispanic/Latino
: num 5 3 7 3 4 2 2 2 1 2 ...
## $ Percent of total enrollment that are Native Hawaiian or Other Pacific
Islander : num 0 0 0 0 0 0 0 0 0 0 ...
## $ Percent of total enrollment that are White
: num 83 62 48 6 75 19 87 62 84 64 ...
## $ Percent of total enrollment that are two or more races
: num 3 1 4 3 2 2 2 2 1 0 ...
## $ Percent of total enrollment that are Race/ethnicity unknown
: num 3 16 3 1 4 4 1 13 10 1 ...
## $ Percent of total enrollment that are Nonresident Alien
: num 0 11 18 3 3 2 3 3 0 3 ...
## $ Percent of total enrollment that are Asian/Native Hawaiian/Pacific
Islander : num 1 4 13 13 3 1 2 3 1 3 ...

```

```
## $ Percent of total enrollment that are women
: num  55 52 49 71 56 71 62 60 68 66 ...
## $ Percent of undergraduate enrollment that are American Indian or Alaska
Native      : num  0 0 1 0 0 1 0 1 0 1 ...
## $ Percent of undergraduate enrollment that are Asian
: num  1 0 16 10 2 1 2 2 1 1 ...
## $ Percent of undergraduate enrollment that are Black or African American
: num  5 0 7 77 10 72 2 15 4 29 ...
## $ Percent of undergraduate enrollment that are Hispanic/Latino
: num  5 0 10 2 4 2 2 2 1 2 ...
## $ Percent of undergraduate enrollment that are Native Hawaiian or Other
Pacific Islander: num  0 0 0 0 0 0 0 0 0 0 ...
## [list output truncated]

dim(collegedata.df)

## [1] 570 130
```

Add a new chunk by clicking the *Insert Chunk* button on the toolbar or by pressing *Ctrl+Alt+I*.

```
collegedata.p <- ggplot(data = collegedata.df, aes_string(x = "`Graduate
enrollment`", y = "`Undergraduate enrollment`", size = "`Graduation rate -
Bachelor degree within 5 years, total`"))
collegedata.p + geom_point(aes_string(color = "`Profile`"), alpha=0.5) +
  labs(title = "Relation between Undergraduate enrollment and Graduate
enrollment ") + scale_radius() + scale_color_manual(values = c("Other" =
"lightsalmon", "Peer" = "yellow", "Aspire" = "red4")) + theme(legend.position
= "bottom")
```

