hw_02

lisa liubovich

2024-01-30

```r
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.4
## v forcats   1.0.0     v stringr   1.5.0
## v ggplot2   3.4.4     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.0
## v purrr     1.0.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become error
```

```r
library(ggplot2)
```

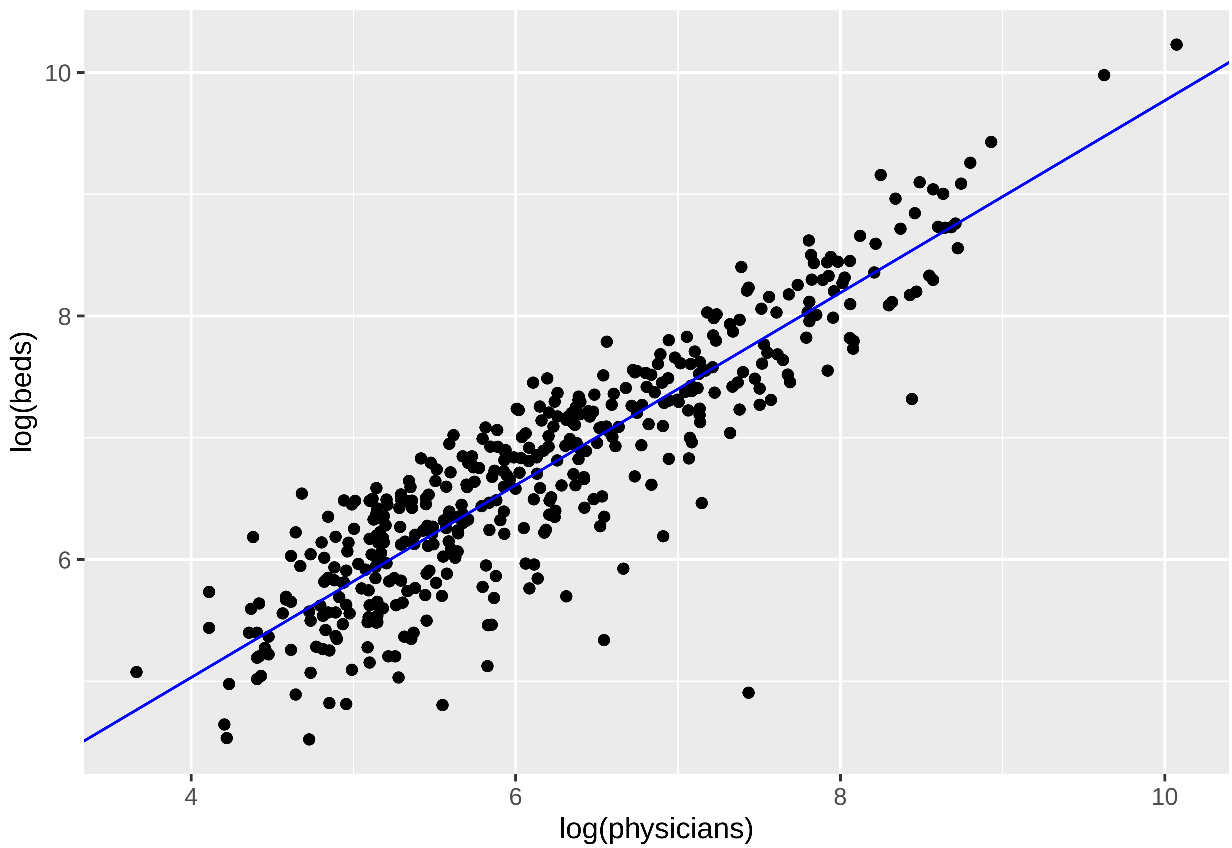
County Demographic Information

1. log(y) = 1.87 + 0.79log(x) n is an example of a power-law relationship. If you exponentiate both sides of the equation, the result is y = $e^{1.87}$ $x^{0.79}$. This follows the form of a power-law relationship, where y = e^{B0} x^{B1}. This form is a simplified version of the original equation form, log(y) = B0 + B1log(x). In this equation, B0 is 1.87 and B1 is 0.79. Therefore, this equation is an example of a power-law relationship.

2.
```r
cdi <- read_csv(file = "../week_2/cdi.csv")
```

```
## Rows: 440 Columns: 17
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (3): county, state, region
## dbl (14): id, area, pop, percent_18_34, percent_65, physicians, beds, crimes...
##
## 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.
```

```r
ggplot(cdi, mapping = aes(x = log(physicians), y = log(beds))) +
         geom_point() +
  geom_abline(intercept = 1.87, slope = 0.79, color = "blue")
```



3. Rewritten/simplified, this equation is $y = e^{1.87} x^{0.79}$.
4. Hospitals with x times as many physicians have $x^{0.79}$ times as many beds, on average; for example, hospitals with twice as many physicians have $2^{0.79}$ as many or 72.91% more beds on average.

```r
2^0.79
```

```
## [1] 1.729074
```

5. $\log(y) = 6.26 + 0.022x$ is an example of an exponential equation. If you were to exponentiate both sides of $\log(y) = B0 + Bx$, the simplified form of the equation is $y = Ae^{B1x}$, where A is B0. $\log(y) = 6.26 + 0.022x$ can be simplified the same way in order to isolate y; if you exponentiate both sides, the resulting equation is $y = 6.26e^{0.022x}$. This is clearly an exponential growth function.
6. x = % of population with a bachelors degree, y = the number of hospital beds. For each additional percentage of people with a bachelors degree, there are $e^{0.022}$ times as many hospital beds, on average; For each additional percentage of people with a bachelors degree, there are 2.2% more hospital beds on average.

```r
exp(0.022)
```

```
## [1] 1.022244
```

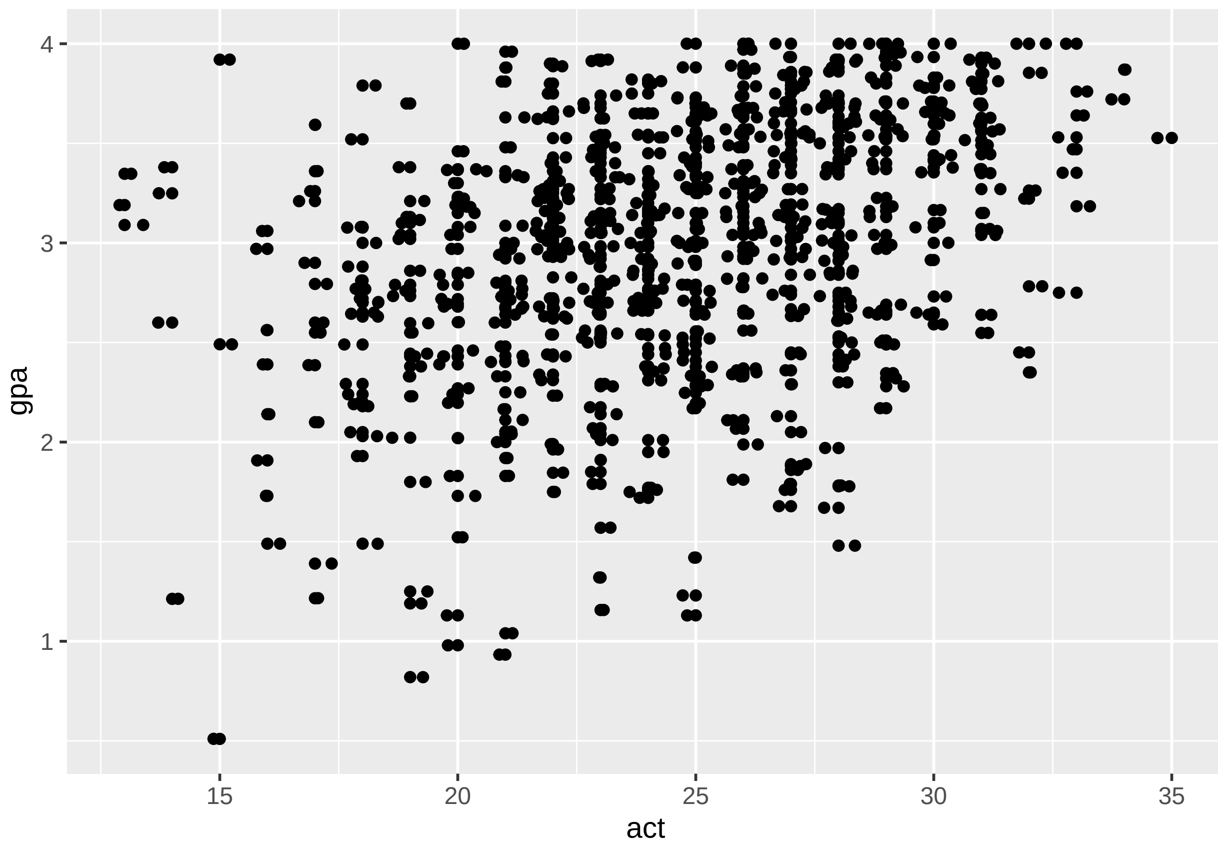
7. 6.26 is the y-intercept, or the number of hospital beds when the percentage of people with a bachelors degree is 0.

University Admissions Data

```r
uni <- read_csv("../week_2/university.csv")
```

```
## Rows: 705 Columns: 5
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## dbl (5): id, gpa, rank, act, year
##
## 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.
```

```r
ggplot(uni, mapping = aes(x = act, y = gpa)) +
  geom_point() +
  geom_jitter()
```



1. There appears to be some sort of moderate positive linear relationship, where as ACT score increases, so does freshman GPA.
2. 1.56 is the y-intercept in this equation. Since the absolute minimum score one can get on the ACT is 1 and not zero, the standard interpretation of a y-intercept for a linear equation isn't appropriate. Instead, we can interpret it like this: The lowest possible score is 1, so y = 1.56 + 0.058(1) –> y = 1.618; therefore, 1.618 is a students freshman GPA if they got the lowest score possible on the ACT. 0.058 is the slope. For every additional unit in ACT score, freshman GPA is 0.058 points higher, on average.

```r
1.56+0.058
```

```
## [1] 1.618
```