SDS 315 HW3

Pranav Rao

2024-02-04

```
## Visit my GitHub repository: [SDS-315](https://github.com/pr10196/SDS-315)

## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.3     v readr     2.1.4
## v forcats   1.0.0     v stringr   1.5.0
## v ggplot2   3.4.3     v tibble    3.2.1
## v lubridate 1.9.2     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 errors

## Warning: package 'mosaic' was built under R version 4.3.2

## Registered S3 method overwritten by 'mosaic':
##   method                           from
##   fortify.SpatialPolygonsDataFrame ggplot2
##
## The 'mosaic' package masks several functions from core packages in order to add
## additional features.  The original behavior of these functions should not be affected by this.
##
## Attaching package: 'mosaic'
##
## The following object is masked from 'package:Matrix':
##
##     mean
##
## The following objects are masked from 'package:dplyr':
##
##     count, do, tally
##
## The following object is masked from 'package:purrr':
##
##     cross
##
## The following object is masked from 'package:ggplot2':
##
##     stat
##
## The following objects are masked from 'package:stats':
##
```

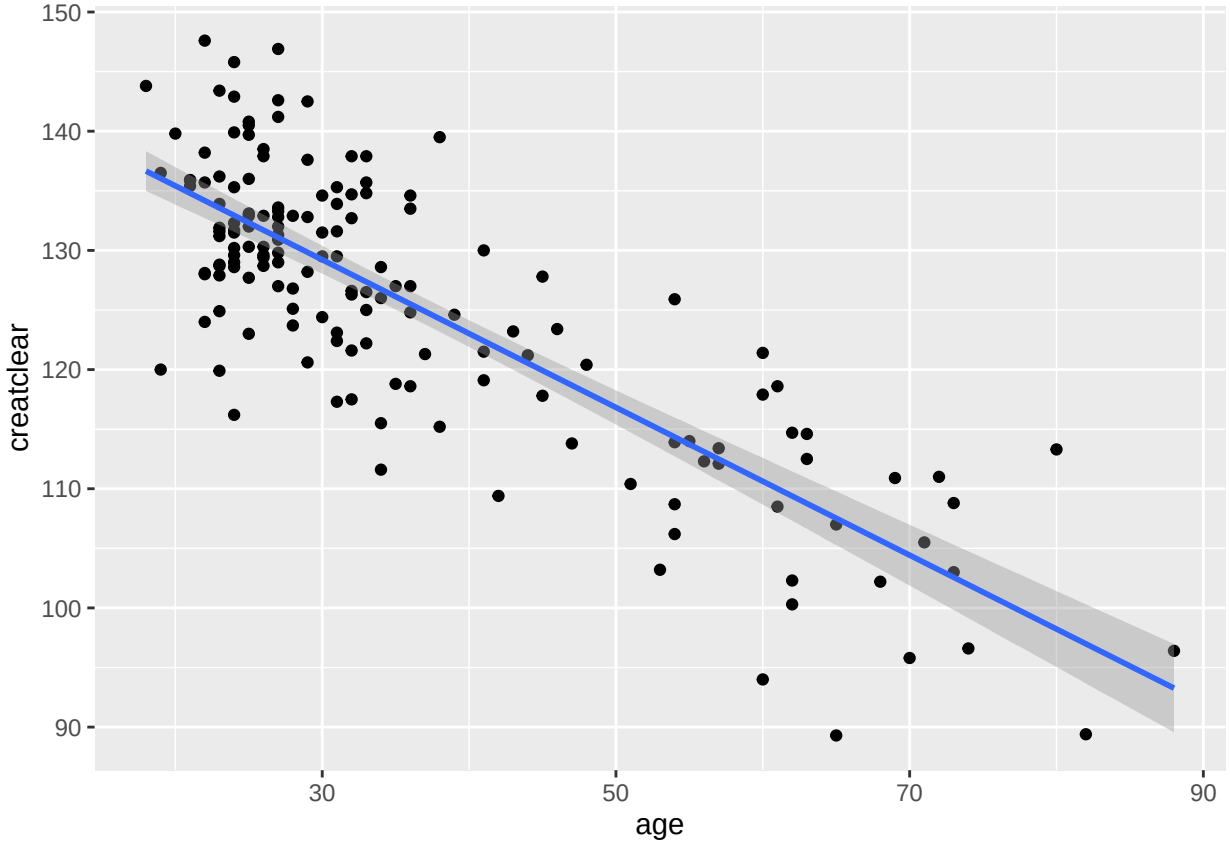
```
##     binom.test, cor, cor.test, cov, fivenum, IQR, median, prop.test,
##     quantile, sd, t.test, var
##
## The following objects are masked from 'package:base':
##
##     max, mean, min, prod, range, sample, sum
```

Problem 1

Part A

```
##   age creatclear
## 1  31      117.3
## 2  36      124.8
## 3  24      145.8
## 4  35      118.8
## 5  53      103.2
## 6  36      127.0
```

```
## `geom_smooth()` using formula = 'y ~ x'
```




```
## (Intercept)         age
## 147.8129158  -0.6198159
```

We should expect a creatinine clearance rate of approximately 113.7. This was determined through the linear regression equation above, 147.8129158-0.6198159x, and the x value for this question was 55, representing the age of a given patient.

Part B

```
## (Intercept)         age
## 147.8129158  -0.6198159
```

On average, the creatinine clearance rate decreases by 0.6198159 mL/minute for every additional year of age. This was determined by calculating the coefficients of the linear regression model and finding the slope of the model, being the "age" variable.

Part C

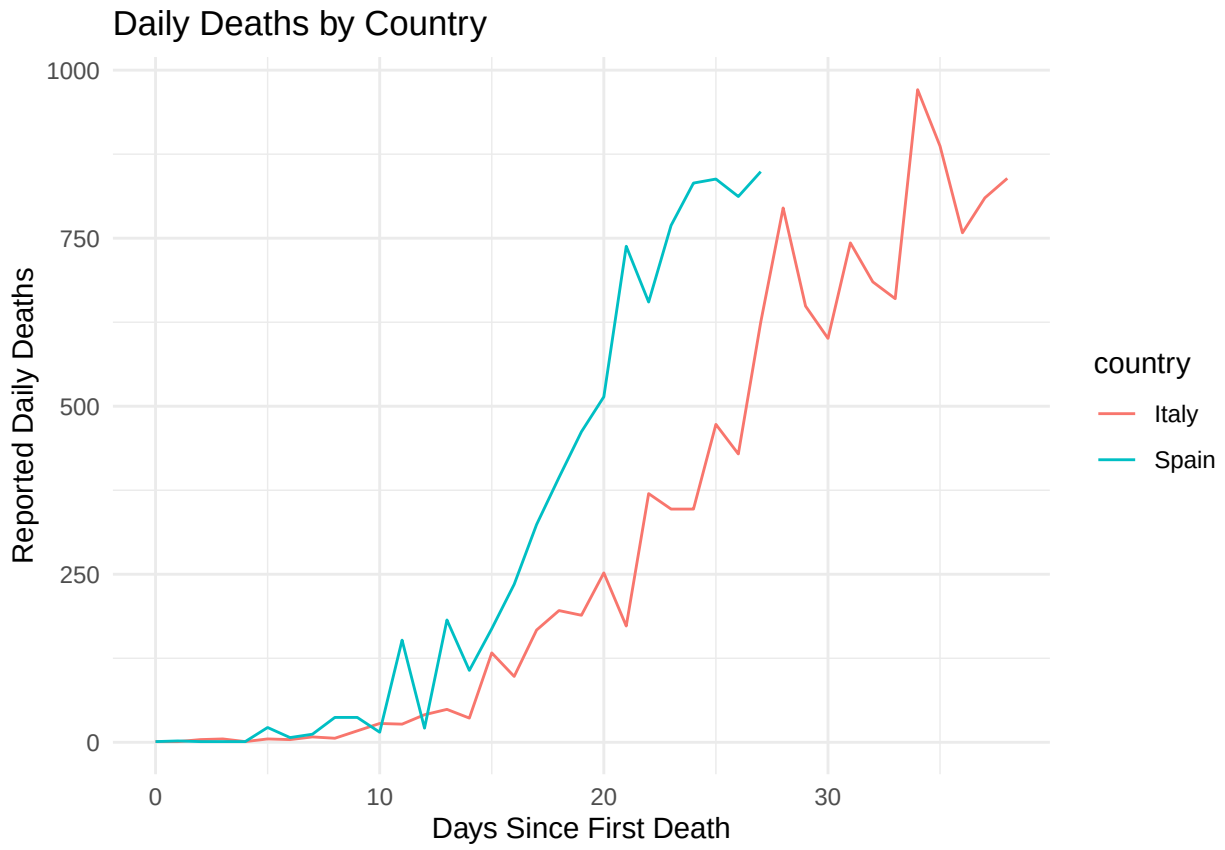
```
## (Intercept)         age
## 147.8129158  -0.6198159
```

The linear regression model equation, 147.8129158-0.6198159(40), calculates a 40-year-old's average rate, which is substituted for x, to be approximately 123, meaning that the rate of 135 is healther than the average rate for a 40-year-old man. In comparison, a 60-year-old man has an average approximate rate of 110.6, according to the equation, 147.8129158-0.6198159(60), which is slightly lower than the 60-year-old man with a rate of 112. Therefore, a 40-year-old with a rate of 135 has a healthier creatinine clearance rate.

Problem 2

The "beta" of a stock measures the systematic risk of a stock and the percentage change of a stock's return when given a one percent chance in the market portfolio. In relation to the linear regression equation, the beta of a stock is calculated by subtracting the B0(k), expected rate of return of stock (k) when the market return is zero, and et(k), the residual for stock "K", from Yt(k), the rate of return of an individual stock "k" and dividing the newly produced difference by Xt, which is the mark rate of return. The table above displays the regression results of the ticker symbols AAPL, GOOG, MRK, JNJ, WMT, and TGT, which include the Intercepts, Slopes, and R_squared values. The GOOG stock displays the strongest correlation, followed by the stocks JNJ, MRK, WMT, TGT, and AAPL. Based on the resulting Slope values produced in the regression results, the WMT stock has the lowest systematic risk, while the AAPL stock has the highest systematic risk.

Problem 3



The estimated growth rate for Italy is 0.183, and the estimated doubling time for Italy is 4 days. In comparison, the estimated growth rate for Italy is 0.276, and the estimated doubling time for Italy is 3 days.

Problem 4

```
##
## Call:
## lm(formula = log_sales ~ log_price, data = milk_dataset)
##
## Coefficients:
## (Intercept)    log_price
##       4.721       -1.619

## log_price
## -1.618578
```

The estimated price elasticity of demand for milk is -1.619, and this estimate was calculated by utilizing the power-law model for the milk sales dataset. Through utilizing this dataset, a regression on the log-log scale was ran and summarized the sales of milk across different prices by calculating the price elasticity of demand for milk, which was represented by the log price variable.