Assignment 10

Daphne Shaw

2022-06-05

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.6     v dplyr   1.0.8
## v tidyr   1.2.0     v stringr 1.4.0
## v readr   2.1.2     v forcats 0.5.1

## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(rvest)
```

```
##
## Attaching package: 'rvest'

## The following object is masked from 'package:readr':
##
##     guess_encoding
```

```
robotstxt::paths_allowed("https://www.nippon.com/en/japan-data/h00603/")
```

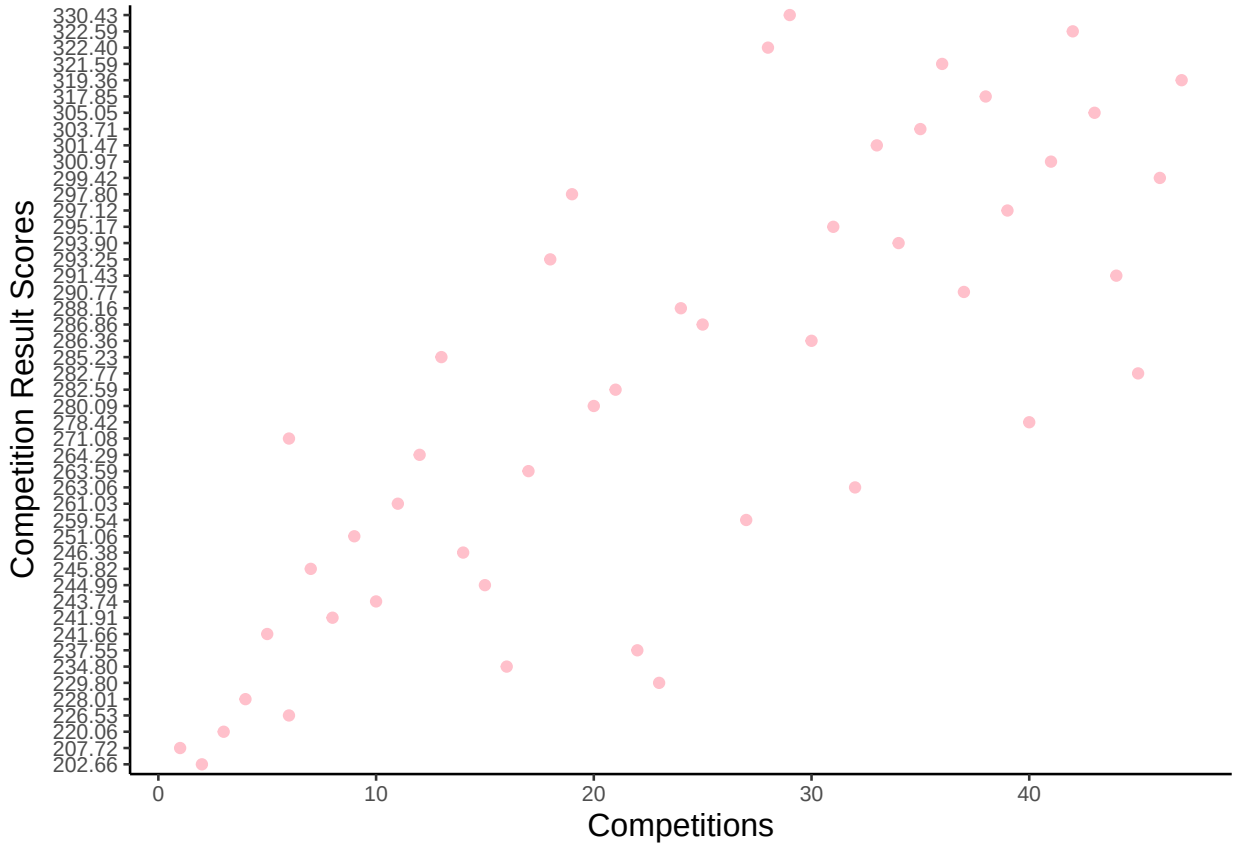
```
## www.nippon.com

##

## [1] TRUE
```

```
page<-read_html("https://www.nippon.com/en/japan-data/h00603/")

scores<-page %>%
  html_nodes(".second_color+ td") %>%
  html_text() %>%
  str_remove("\\(") %>%
  str_remove("\\)") %>%
  str_remove("\\Does not compete") %>%
```

```
  str_remove("[st]") %>%
  str_remove("[nd]") %>%
  str_remove("[rd]") %>%
  str_remove("[t]") %>%
  str_remove("[h]")


yuzuru<-tibble(
  scores=scores
)

yuzuru2 <- yuzuru[-c(1,3,4,11,15,19), ]

yuzuru2$scores<-substring(yuzuru2$scores,2)

comp0<-c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,6,27,28,29,30,31,32,33,34,35
         36,37,38,39,40,41,42,43,44,45,46,47)
Competitions<-rev(comp0)

yuzuru3<-
  yuzuru2 %>%
  summarise(Competitions,scores)

print(yuzuru3)
```

```
## # A tibble: 47 x 2
##    Competitions scores
##           <dbl> <chr>
##  1           47 " 319.36"
##  2           46 " 299.42"
##  3           45 " 282.77"
##  4           44 " 291.43"
##  5           43 " 305.05"
##  6           42 " 322.59"
##  7           41 " 300.97"
##  8           40 " 278.42"
##  9           39 " 297.12"
## 10           38 " 317.85"
## # ... with 37 more rows
```

```
ggplot(yuzuru3,aes(x=Competitions,y=scores))+
  geom_point(color='pink')+
  xlab("Competitions")+
  ylab("Competition Result Scores")+
  theme_classic(base_size=10)+
  theme(axis.title=element_text(size=12))
```



The graph above uses data from www.nippon.com, and is a scatter plot of the figure skater, Yuzuru Hanyu's figrue skating competition scores over various competitions, beginning from the Grand Prix NHK Trophy from the 2010-2011 season. The plot above shows Yuzuru Hanyu's competition result scores as a function of the number of competitions he has been to over the years, and shows an approximately positive correlation, meaning that as the amount of competitions he has been to increases, so does his score. This likely indicates a positive linear regression. In conclusion, according to the data from www.nippon.com and the subsequent graph drawn above, we can conclude that from this data alone, as the amount of competitions Yuzuru Hanyu attends increases, so does his resulting skate score.