R Notebook

R Studio API Code

```
knitr::opts_chunk$set(echo = TRUE)
# library(rstudioapi)
# setwd(dirname(rstudioapi::getActiveDocumentContext()$path))
```

Libraries

```
library(tidyverse)
```

```
## -- Attaching packages ----------------------------------------------------------
```

```
## v ggplot2 3.2.1      v purrr   0.3.3
## v tibble  2.1.3      v dplyr   0.8.4
## v tidyr   1.0.2      v stringr 1.4.0
## v readr   1.3.1      v forcats 0.5.0
```

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

```
library(httr)
library(rvest)
```

```
## Loading required package: xml2
```

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

```
## The following object is masked from 'package:purrr':
##
##     pluck
```

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

Data Import and Cleaning

```
profile <- read_html("https://scholar.google.com/citations?user=nx_vfuYAAAAJ&hl=en&oi=ao")
profile_nodes <- html_nodes(profile,"#gsc_a_b .gs_ibl , .gsc_a_at+ .gs_gray , .gsc_a_at")
profile_text <- html_text(profile_nodes)
profile_tbl <- as_tibble(matrix(profile_text,nrow=20,ncol=4,byrow=TRUE,
                                dimnames=list(c(),
```

```
                                              c("article","authors","citationCount","year")))) %>%
  mutate(year=as.numeric(year),
         citationCount=as.numeric(citationCount))
profile_tbl
```

```
## # A tibble: 20 x 4
##    article                      authors             citationCount  year
##    <chr>                        <chr>                       <dbl> <dbl>
##  1 The relative importance of task, ~ M Rotundo, PR Sackett      1635  2002
##  2 Interpersonal deviance, organizat~ CM Berry, DS Ones, PR~     1336  2007
##  3 Research strategies and tactics i~ PR Sackett, JR Larson~      790  1990
##  4 Investigating the dimensionality ~ ML Gruys, PR Sackett        760  2003
##  5 Counterproductive Behavior at Work PR Sackett                  748  2001
##  6 The structure of counterproductiv~ PR Sackett                  566  2002
##  7 High-stakes testing in employment~ PR Sackett, N Schmitt~      547  2001
##  8 Social desirability corrections i~ JE Ellingson, PR Sack~      455  1999
##  9 Constructs and assessment center ~ PR Sackett, GF Dreher       429  1982
## 10 Relations between measures of typ~ PR Sackett, S Zedeck,~      427  1988
## 11 A meta-analytic review of gender ~ M Rotundo, DH Nguyen,~      423  2001
## 12 Personnel selection                PR Sackett, F Lievens       416  2008
## 13 Integrity testing for personnel s~ PR Sackett, LR Burris~      397  1989
## 14 Tokenism in performance evaluatio~ PR Sackett, CL DuBois~      372  1991
## 15 Within-group norming and other fo~ PR Sackett, SL Wilk         361  1994
## 16 Factors related to the satisfacti~ JE Ellingson, ML Gruy~      359  1998
## 17 New developments in the use of me~ PR Sackett, JE Wanek        356  1996
## 18 Correction for range restriction:~ PR Sackett, H Yang          352  2000
## 19 Citizenship and counterproductive~ PR Sackett, CM Berry,~      310  2006
## 20 How are we doing after 30 years? ~ F Anseel, AS Beatty, ~      307  2015
```

Analysis

The correlation beween year and citation count is 0.1836999, with a p-value of 0.4381826.

Visualization

The Relationship between Year and Citation Count of Paul Sackett's Public