Rise and Fall of Programming Languages

Thu Rein Aung

2024-09-05

Load necessary libraries

```
library(readr)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(ggplot2)
```

Load and Inspecting the data

```
programming_data <- read.csv("programming_data.csv")
head(programming_data)
```

```
##   year           tag number year_total
## 1 2008     .htaccess     54      58390
## 2 2008          .net   5910      58390
## 3 2008      .net-2.0    289      58390
## 4 2008      .net-3.5    319      58390
## 5 2008      .net-4.0      6      58390
## 6 2008 .net-assembly      3      58390
```

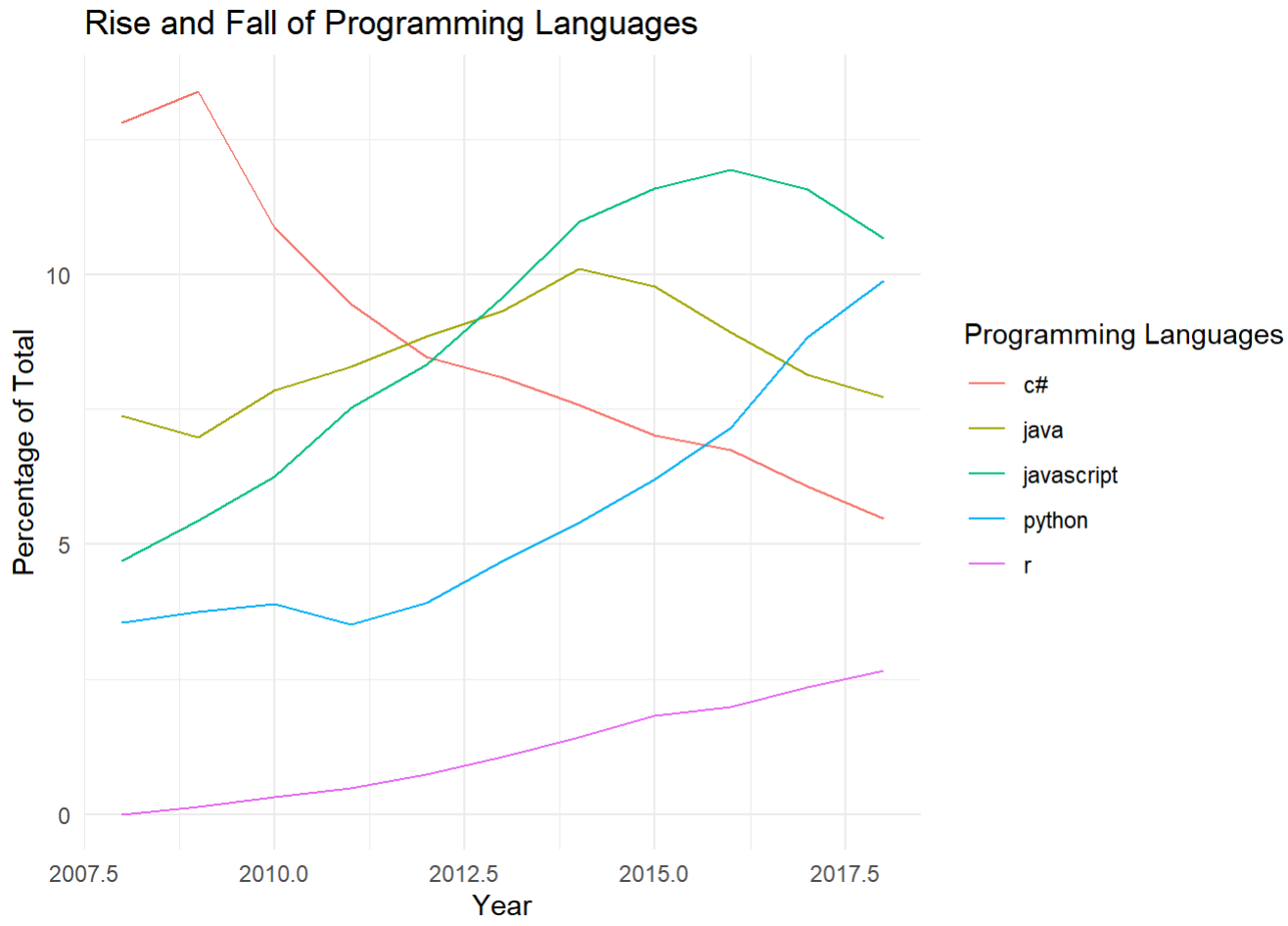
1.Trend Analysis: Percentage share of

programming languages over the years

```
all_programming_data <- programming_data %>%
  mutate(percentage = round((number / year_total) * 100, 2))

# Visualizing trends for popular programming languages
popular_languages <- c("c#", "java", "javascript", "python", "r")
selected_data <- all_programming_data %>%
  filter(tag %in% popular_languages)

ggplot(selected_data, aes(x = year, y = percentage, color = tag)) +
  geom_line() +
  labs(title = "Rise and Fall of Programming Languages",
       x = "Year",
       y = "Percentage of Total",
       color = "Programming Languages") +
  theme_minimal()
```



2. Most Popular Programming Languages Each Year

```
most_popular <- all_programming_data %>%
  group_by(year) %>%
  top_n(1, number)

cat("Most Popular Language Each Year:\n")
```

```
## Most Popular Language Each Year:
```

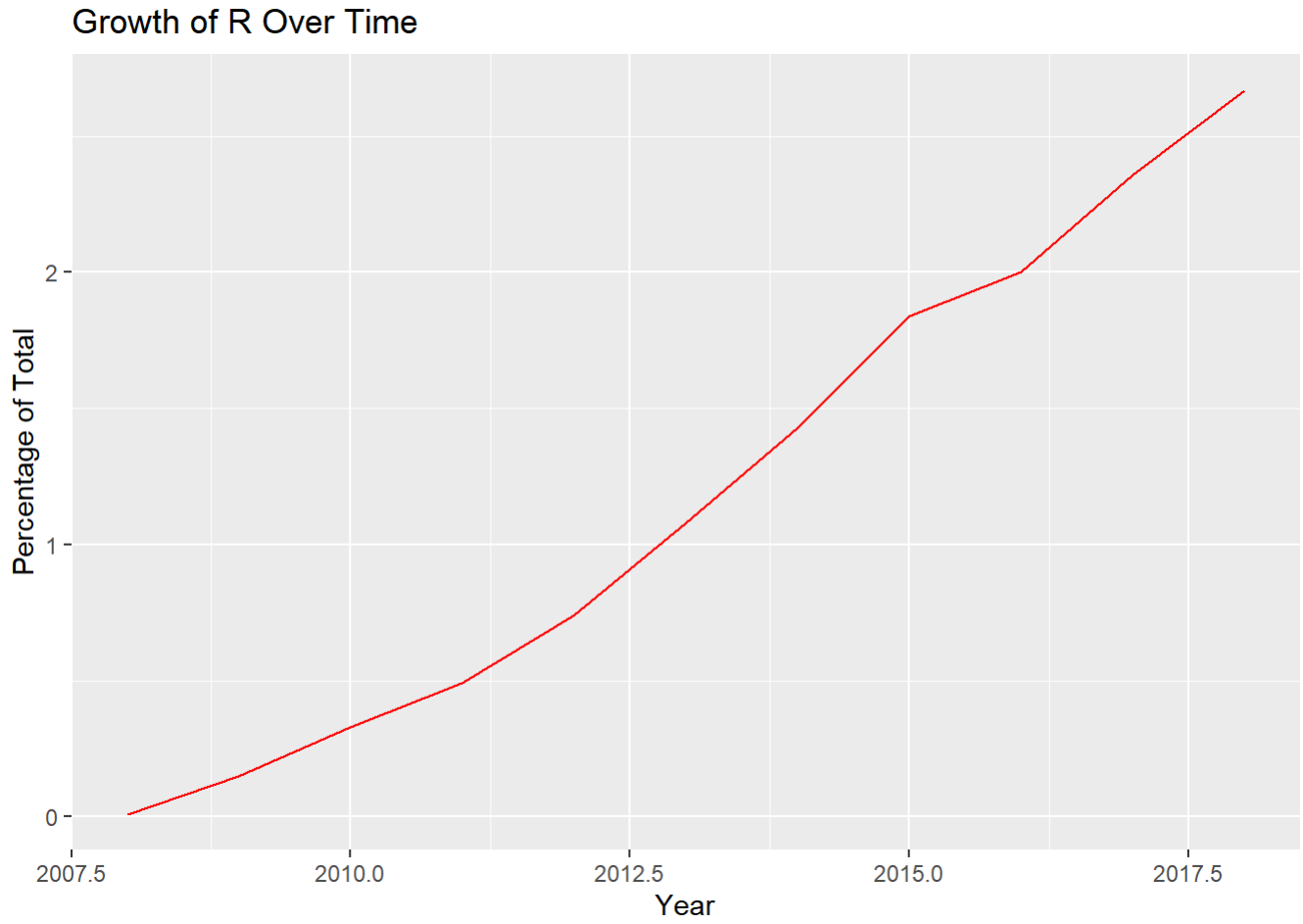
```
print(most_popular)
```

```
## # A tibble: 11 × 5
## # Groups:   year [11]
##     year tag        number year_total percentage
##    <int> <chr>       <int>      <int>      <dbl>
##  1  2008 c#           7473      58390      12.8
##  2  2009 c#          46044     343868      13.4
##  3  2010 c#          75501     694391      10.9
##  4  2011 c#         113408    1200551       9.45
##  5  2012 java       145640    1645404       8.85
##  6  2013 javascript 197101    2060473       9.57
##  7  2014 javascript 237415    2164701      11.0
##  8  2015 javascript 257006    2219527      11.6
##  9  2016 javascript 265896    2226072      11.9
## 10  2017 javascript 266762    2305207      11.6
## 11  2018 javascript 115726    1085170      10.7
```

3. Has R been growing or shrinking over time?

```
r_over_time <- all_programming_data %>%
  filter(tag == "r")

ggplot(r_over_time, aes(x = year, y = percentage)) +
  geom_line(color = "red") +
  labs(title = "Growth of R Over Time",
       x = "Year",
       y = "Percentage of Total")
```



4. Emerging and Declining Languages

```
# Calculate the percentage change in tag occurrences year-over-year
emerging_declining_languages <- all_programming_data %>%
  group_by(tag) %>%
  arrange(year) %>%
  mutate(perc_change = round((number - lag(number)) / lag(number) * 100, 2))

# Filter languages with significant positive or negative changes
emerging_languages <- emerging_declining_languages %>%
  filter(perc_change > 50)

declining_languages <- emerging_declining_languages %>%
  filter(perc_change < -50)

cat("Emerging Languages (More than 50% growth in any year):\n")
```

```
## Emerging Languages (More than 50% growth in any year):
```

```
print(emerging_languages)
```

```
## # A tibble: 11,599 × 6
## # Groups:   tag [4,068]
##     year tag           number year_total percentage perc_change
##    <int> <chr>          <int>      <int>      <dbl>       <dbl>
##  1  2009 .htaccess        828     343868       0.24       1433.
##  2  2009 .net           23076     343868       6.71        290.
##  3  2009 .net-2.0         593     343868       0.17        105.
##  4  2009 .net-3.5        1087     343868       0.32        241.
##  5  2009 .net-4.0         129     343868       0.04       2050
##  6  2009 .net-assembly     13     343868       0           333.
##  7  2009 2d               143     343868       0.04        240.
##  8  2009 32-bit            99     343868       0.03        421.
##  9  2009 32bit-64bit       63     343868       0.02       1475
## 10  2009 3d               414     343868       0.12        467.
## # ℹ 11,589 more rows
```

```
cat("Declining Languages (More than 50% decline in any year):\n")
```

```
## Declining Languages (More than 50% decline in any year):
```

```
print(declining_languages)
```

```
## # A tibble: 3,291 × 6
## # Groups:   tag [2,796]
##     year tag                       number year_total percentage perc_change
##    <int> <chr>                      <int>      <int>      <dbl>       <dbl>
##  1  2009 template-meta-programming      1     343868       0          -75
##  2  2010 asp.net-web-api                3     694391       0          -57.1
##  3  2010 child-process                  2     694391       0          -75
##  4  2010 ssrs-2008-r2                   1     694391       0          -80
##  5  2010 subsonic                     304     694391       0.04       -55.8
##  6  2010 windows-vista                312     694391       0.04       -52.9
##  7  2011 agile                         92    1200551       0.01       -51.6
##  8  2011 google-sheets-api              1    1200551       0          -66.7
##  9  2011 latex                        405    1200551       0.03       -61.0
## 10  2011 maven-2                      954    1200551       0.08       -50.6
## # ℹ 3,281 more rows
```

5. Peak Popularity: Finding the year when each language was most popular

```
peak_popularity <- all_programming_data %>%
  group_by(tag) %>%
  filter(number == max(number))

cat("Peak Popularity of Each Language:\n")
```

```
## Peak Popularity of Each Language:
```

```
print(peak_popularity)
```

```
## # A tibble: 4,122 × 5
## # Groups:   tag [4,080]
##     year tag                number year_total percentage
##    <int> <chr>               <int>      <int>      <dbl>
##  1  2009 .net-2.0              593     343868       0.17
##  2  2009 agile                 202     343868       0.06
##  3  2009 build-process         469     343868       0.14
##  4  2009 compact-framework     657     343868       0.19
##  5  2009 nant                  220     343868       0.06
##  6  2009 project-management    401     343868       0.12
##  7  2009 remoting              220     343868       0.06
##  8  2009 subsonic              687     343868       0.2
##  9  2009 visual-studio-2005    801     343868       0.23
## 10  2009 windows-mobile        805     343868       0.23
## # ℹ 4,112 more rows
```