one-worded-vs-multi-worded

Here I looked at how company’s name effects its popularity on LinkedIn. I found that companies with a single word in their name have about 4 times more followers on average. Therefore, I suggest that companies with names like e.g. “Liberty Tax Service” use abbreviation: “LTS”; Companies with names like “Ashland Inc” should simply use Ashland.

```
linkedin<-read.csv("C:\\Users\\irakl\\Desktop\\temp_datalab_records_linkedin_company\\linkedin.c
sv", header = TRUE)
```

```
# work with most recent data
library(dplyr)
```

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

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

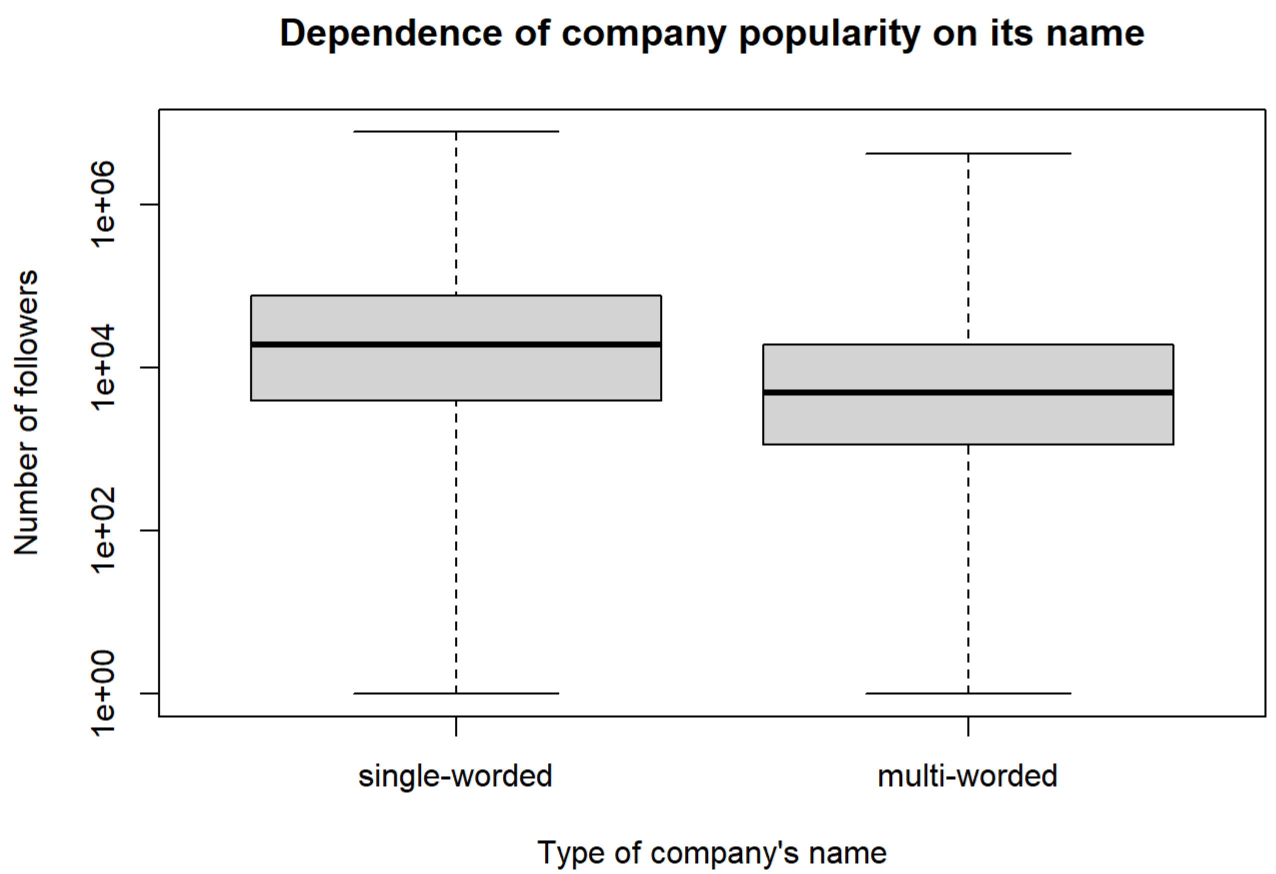
```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
linkedin_recent <-
  linkedin %>%
  group_by(company_name) %>%
  filter(as_of_date == max(as_of_date))
```

```
# produce a column that indicates whether company name is one-worded or multi-worded.
linkedin_recent %>%
  select(company_name,followers_count,industry) %>%
  mutate(
    multi_worded_name = sapply(strsplit(company_name, " "), length) != 1 # e.g. Hewlett-Packard
 is considered one-worded with this code
  ) ->
  name_role
```

Following analysis shows that companies with single-worded name tend to attract few hundred percent more followers.

```
name_role_plus <- name_role
name_role_plus$followers_count <- name_role_plus$followers_count + 1 # added 1 follower to every
one to plot on log scale
boxplot(followers_count ~ multi_worded_name, data = name_role_plus, log = "y", range = 0, names=
c("single-worded","multi-worded"), xlab = "Type of company's name",
   ylab = "Number of followers", main = "Dependence of company popularity on its name")
```




```
t.test(followers_count~multi_worded_name,data = name_role_plus)
```

```
##
##  Welch Two Sample t-test
##
## data:  followers_count by multi_worded_name
## t = 7.4801, df = 1259.9, p-value = 1.387e-13
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   79354.43 135778.24
## sample estimates:
## mean in group FALSE  mean in group TRUE
##           143259.80            35693.47
```

```
# produce follower's vector for companies with multi-worded names
name_role %>%
  filter(multi_worded_name == TRUE) %>%
  ungroup() %>%
  select(followers_count) ->
  multi_worded
colnames(multi_worded) <- NULL

# produce follower's vector for companies with single-worded names
name_role %>%
  filter(multi_worded_name == FALSE) %>%
  ungroup() %>%
  select(followers_count) ->
  single_worded
colnames(single_worded) <- NULL
```

```
ratio <- sapply(single_worded, mean, na.rm = TRUE)/sapply(multi_worded, mean, na.rm = TRUE)
rt = toString(round(ratio))
paste("Companies with names consisting of a single word have about",rt," times more followers")
```

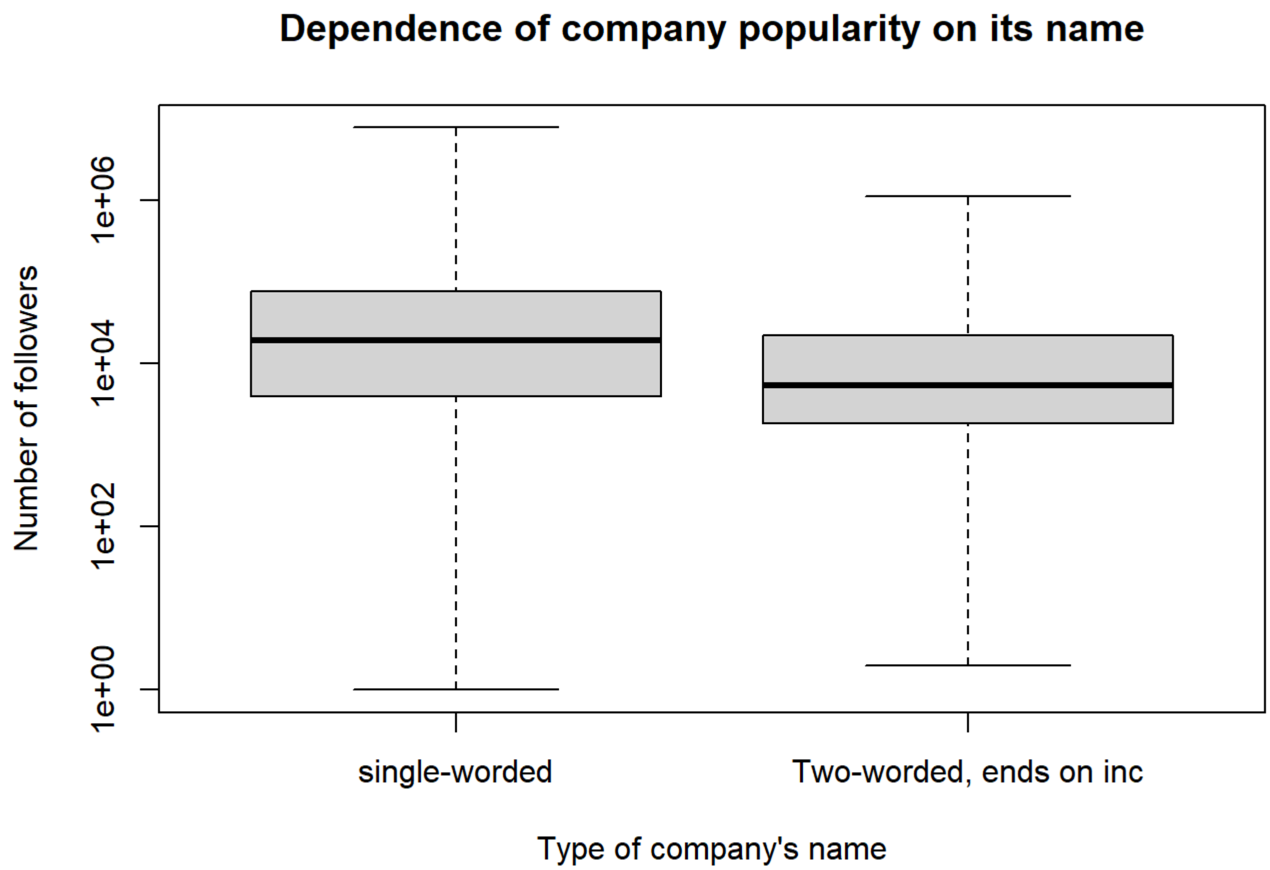
```
## [1] "Companies with names consisting of a single word have about 4  times more followers"
```

```
# produce a column that indicates whether company's name is one-worded or one word +Inc./Inc
#Other companies are neglected
linkedin_recent %>%
  select(company_name,followers_count,industry) %>%
  filter(sapply(strsplit(company_name, " "), length) == 1 |
        (sapply(strsplit(company_name, " "), length) == 2)& grepl(" Inc$| Inc.$", company_name))
%>%
   mutate(
    ends_with_inc = sapply(strsplit(company_name, " "), length) != 1
  ) ->
  inc_role
```

```
View(inc_role)
```

Following analysis clearly demonstrates the importance of one simple change companies can make to increase their number of followers on LinkedIn: removing Inc at the end of their name. For example, GGP Inc. would become GGP.

```
inc_role_plus <- inc_role
inc_role_plus$followers_count <- inc_role_plus$followers_count + 1 # added 1 follower to everyon
e to plot on log scale
boxplot(followers_count ~ ends_with_inc, data = inc_role_plus, log = "y", range = 0, names=c("si
ngle-worded","Two-worded, ends on inc"), xlab = "Type of company's name",
   ylab = "Number of followers", main = "Dependence of company popularity on its name")
```




```
t.test(followers_count~ends_with_inc,data = inc_role_plus)
```

```
##
##  Welch Two Sample t-test
##
## data:  followers_count by ends_with_inc
## t = 6.5527, df = 1310, p-value = 8.102e-11
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   74222.42 137655.38
## sample estimates:
## mean in group FALSE  mean in group TRUE
##            143259.8             37320.9
```