Popularity across industries

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

# extra
library(dplyr)</pre>
```

```
##
## 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 %>%
  group_by(industry) %>%
  summarise_at(vars(followers_count),funs(mean(.,na.rm=TRUE))) %>%
  rename(
    mean_followers = followers_count,
    ) ->
  fol_per_ind
```

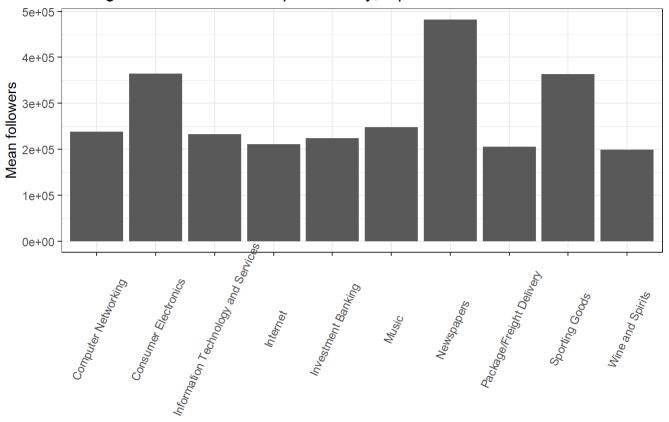
```
## Warning: `funs()` is deprecated as of dplyr 0.8.0.
## Please use a list of either functions or lambdas:
##
##
     # Simple named list:
##
     list(mean = mean, median = median)
##
     # Auto named with `tibble::lst()`:
##
     tibble::lst(mean, median)
##
##
##
     # Using lambdas
     list(~ mean(., trim = .2), ~ median(., na.rm = TRUE))
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_warnings()` to see where this warning was generated.
```

```
fol_per_ind$industry[1]="Other" # Renaming blank
```

```
fol_per_ind_sorted <- fol_per_ind[order(-fol_per_ind$mean_followers),]
View(fol_per_ind_sorted)</pre>
```

It appears that different industries attract a different number of people. However, one should carefully look at outliers and spread in the number of followers in each industry before jumping to conclusions.

Average number of followers per Industry, top ten



Industry