Myproject.R

Admin

2020-03-12

library(readxl)  
library(tidyverse)

## -- Attaching packages ----------------------------------------------------------------- tidyverse 1.3.0 --

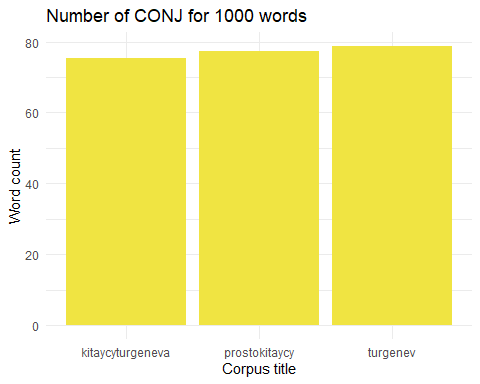
## v ggplot2 3.2.1 v purrr 0.3.3  
## v tibble 2.1.3 v dplyr 0.8.3  
## v tidyr 1.0.2 v stringr 1.4.0  
## v readr 1.3.1 v forcats 0.4.0

## -- Conflicts -------------------------------------------------------------------- tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

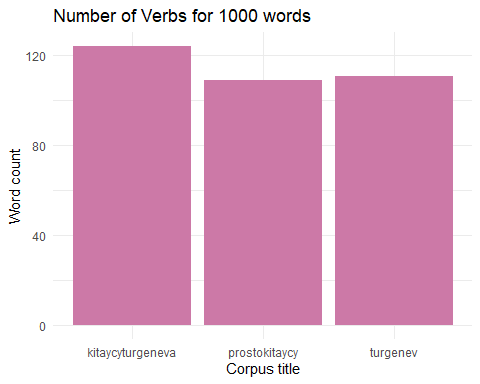
x <- read\_excel("kitaycy.xlsx")

## New names:  
## \* `` -> ...1  
## \* `` -> ...2  
## \* `` -> ...3  
## \* `` -> ...4

x <- x%>%  
 fill(all)  
x1 <- x %>%  
 rename(POS = ...3)%>%  
 rename(corpusname = ...1)%>%  
 rename(POScount = ...4)%>%  
 rename(corp\_wds\_count = all)  
x1 <- x1%>%  
 mutate(POScount\_for\_kw = (POScount/corp\_wds\_count)\*1000)   
x1 <- select (x1,-c(...2))  
x1 <- x1%>%  
 arrange(POS)  
nr\_of\_adj <- x1%>%  
 filter(POS == "ADJF")  
nr\_of\_Verbs <- x1%>%  
 filter(POS == "VERB")  
nr\_of\_conj <- x1%>%  
 filter(POS == "CONJ")  
nr\_of\_conj %>%  
 arrange(POScount\_for\_kw) %>%  
 ggplot(., aes(x= corpusname, y=POScount\_for\_kw)) +  
 geom\_bar(stat='identity', fill="#F0E442") +   
 ylab("Word count") + xlab("Corpus title") +  
 ggtitle("Number of CONJ for 1000 words") +   
 theme\_minimal()



nr\_of\_Verbs %>%  
 arrange(POScount\_for\_kw) %>%  
 ggplot(., aes(x= corpusname, y=POScount\_for\_kw)) +  
 geom\_bar(stat='identity', fill="#CC79A7") +   
 ylab("Word count") + xlab("Corpus title") +  
 ggtitle("Number of Verbs for 1000 words") +   
 theme\_minimal()



nr\_of\_adj %>%  
 arrange(POScount\_for\_kw) %>%  
 ggplot(., aes(x= corpusname, y=POScount\_for\_kw)) +  
 geom\_bar(stat='identity', fill="#1CCCC6") +   
 ylab("Word count") + xlab("Corpus title") +  
 ggtitle("Number of adjectives for 1000 words") +   
 theme\_minimal()

