

president_countries

Loading library

```
library("tidyverse")

## -- Attaching packages ----

## v ggplot2 3.3.0      v purrr   0.3.3
## v tibble   2.1.3      v dplyr    0.8.5
## v tidyverse 1.0.2     v stringr  1.4.0
## v readr    1.3.1      vforcats  0.4.0

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

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
```

Loading data

we'll scrap data from wikipedia

```
webpage <- read_html("https://en.wikipedia.org/wiki/List_of_state_leaders_in_2019#Africa")
```

Cleaning data

```

texte <- webpage %>%
  html_nodes("a , li+ li li")%>%
  html_text()

texte_1 <- texte[str_detect(texte,"present")]

texte_2 <- texte_1[str_detect(texte_1,"Minister",negate = TRUE)]

df <- data.frame(president = texte_2)

data_president <- df %>%
  separate(president,c("president", "country"),sep = ",")%>%
  filter(str_detect(country,"present"))%>%
  mutate(pays = str_replace_all(country,".*of",""),
         pays = str_replace_all(pays,"\\(.*)",""),
         annee_debut = str_extract_all(country,"\\d\\d\\d\\d"),
         present = 2020,
         president = str_replace_all(president,".*-",""))%>%
  filter(str_detect(annee_debut,"[[:alpha:]]",negate = TRUE))%>%
  mutate(annee_debut = as.numeric(annee_debut),
         annee_au_pouvoir = present - annee_debut)%>%
  distinct(president,.keep_all = TRUE)

data_president[159,1] <- "Simonetta Sommaruga"

```

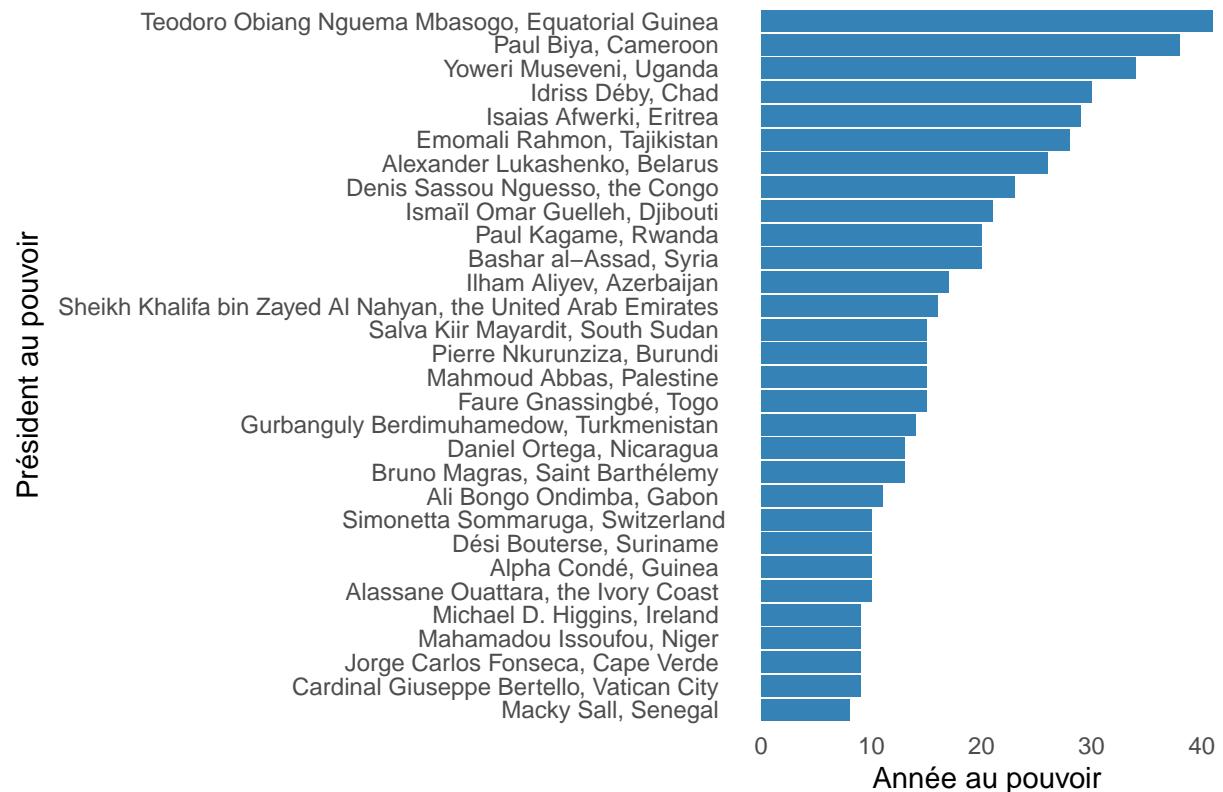
Who are the most president who stayed longer in power ?

```

data_president%>%
  filter(str_detect(country,"President"))%>%
  mutate(president_country = paste0(president,",",pays),
         president_country = fct_reorder(president_country,annee_au_pouvoir))%>%
  arrange(desc(annee_au_pouvoir))%>%
  head(30)%>%
  ggplot(mapping = aes(president_country,annee_au_pouvoir))+
  geom_col(fill = "#3582b6")+
  coord_flip()+
  theme_minimal()+
  theme(panel.grid = element_blank(),
        plot.title = element_text(hjust = 0.08))+
  labs(y = "Année au pouvoir", x = "Président au pouvoir", title = "Les 30 présidents qui sont au pouvoi")

```

Les 30 présidents qui sont au pouvoir



How about looking at whole continent ? Which continent has the most president that had more than two mandate ?

```
data_president_continent <- data_president %>%
  mutate(pays = str_trim(pays))%>%
  inner_join(countries, by =c("pays"="country"))
```

Graph

```
data_president_continent %>%
  filter(str_detect(country,"President"))%>%
  filter(annee_au_pouvoir >= 10)%>%
  group_by(continent)%>%
  summarise(total = n())%>%
  mutate(continent = fct_reorder(continent,total))%>%
  ggplot(mapping = aes(continent , total , fill = continent ))+
  geom_col()+
  coord_flip()+
  theme_minimal()+
  theme(panel.grid = element_blank(),
        plot.title = element_text(hjust = 0.08))+
```

```
scale_y_continuous(breaks = 0:15)+  
  labs(x = "Continent",y="Nombre d'année au pouvoir",title = "Président qui sont au pouvoir depuis au moins dix ans par continent")
```

