---

title: “R Notebook”

output: html_notebook

This is an R Markdown (http://rmarkdown.rstudio.com) Notebook. When you execute code within the notebook, the results appear beneath the code.

Try executing this chunk by clicking the *Run* button within the chunk or by placing your cursor inside it and pressing *Cmd+Shift+Enter*.

```
{r} # plot(cars) #
```

Add a new chunk by clicking the *Insert Chunk* button on the toolbar or by pressing *Cmd+Option+I*.

When you save the notebook, an HTML file containing the code and output will be saved alongside it (click the

Preview button or press *Cmd+Shift+K* to preview the HTML file).

The preview shows you a rendered HTML copy of the contents of the editor. Consequently, unlike *Knit*, *Preview* does not run any R code chunks. Instead, the output of the chunk when it was last run in the editor is displayed.

```
#load tidyverse
library(tidyverse)
```

```
## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──
## ✔ dplyr     1.1.4     ✔ readr     2.1.5
## ✔ forcats   1.0.0     ✔ stringr   1.5.1
## ✔ ggplot2   3.5.1     ✔ tibble    3.2.1
## ✔ lubridate 1.9.3     ✔ tidyr     1.3.1
## ✔ purrr     1.0.2
## ── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
setwd('~/Downloads/Account ownership at a financial institution or with a mobile-money-service provider, female
(% of population ages 15+).')
female <- read.csv("42656245-233e-42cb-a412-d8a2fe924acb_Data.csv")
setwd('~/Downloads/Account ownership at a financial institution or with a mobile-money-service provider, male (%
of population ages 15+).')
male <- read.csv("4fa40f30-e794-41b2-9a30-b425e363385c_Data.csv")

# Determine that R does consider female to be data frame
is.data.frame(female)
```

```
## [1] TRUE
```

```
# Determine that R does consider female to be data frame
is.data.frame(male)
```

```
## [1] TRUE
```

```
library(tidyverse)

rel_col <- which(colnames(female)=="Country.Name" | colnames(female)== "X2014..YR2014." | colnames(female)== "X20
17..YR2017." | colnames(female) == "X2021..YR2021." )
rel_col
```

```
## [1]  3  9 12 16
```

```
#filters specific countries
female <- female %>%
  filter(Country.Name == "Brazil" | Country.Name == "India" | Country.Name == "United States")
female <- female[rel_col] %>% rename("2014" = 2, "2017" = 3, "2021" = 4)

#filters specific years
male <- male %>%
  filter(Country.Name == "Brazil" | Country.Name == "India" | Country.Name == "United States")
male <- male[rel_col] %>% rename("2014" = 2, "2017" = 3, "2021" = 4)

#swaps axes
transpose_f <- data.frame(t(female[-1]))
colnames(transpose_f) <- female[, 1]
transpose_m <- data.frame(t(male[-1]))
colnames(transpose_m) <- male[, 1]

#prints data type
print(sapply(transpose_f, class))
```

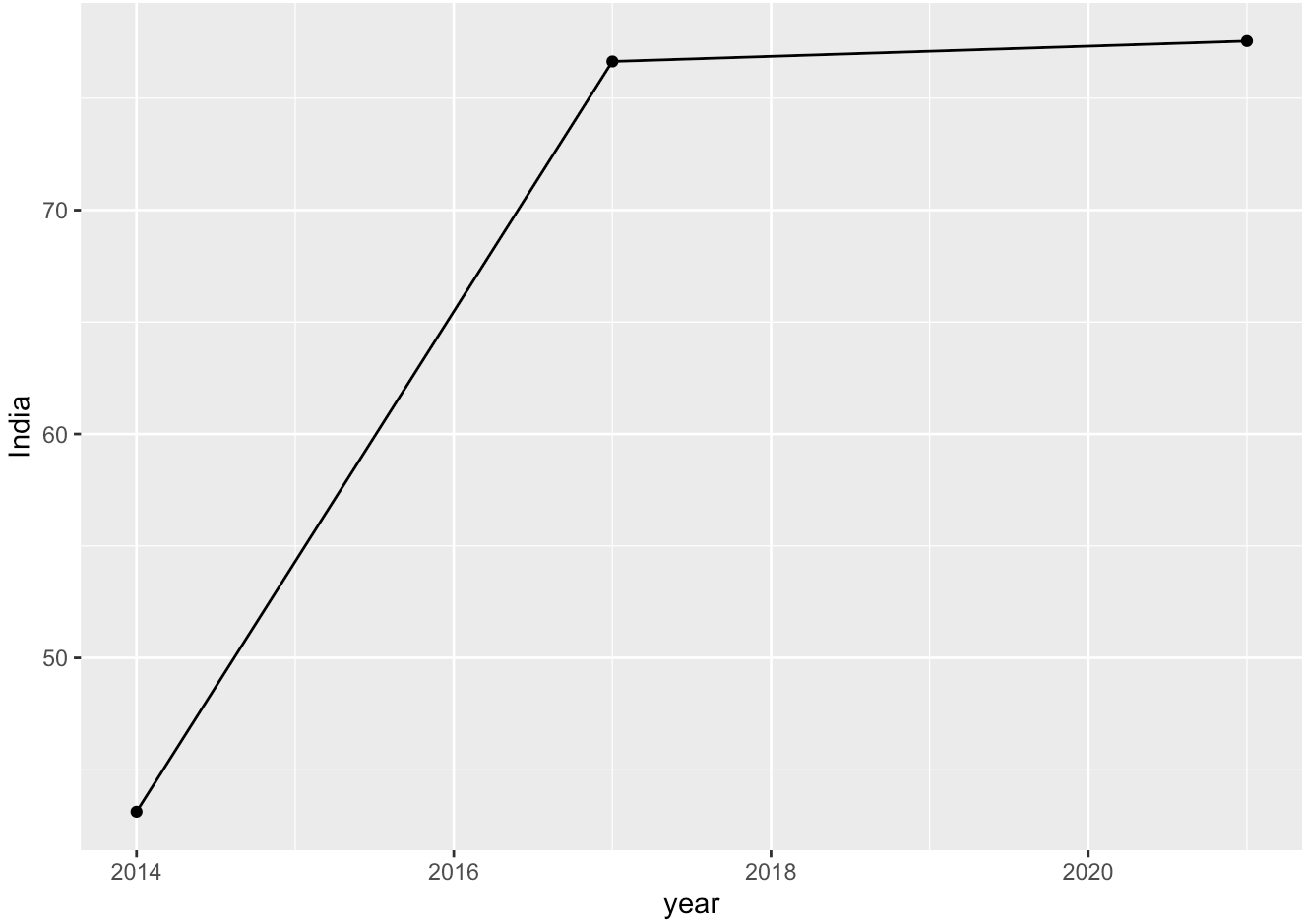
```
##        Brazil         India United States
##   "character"   "character"   "character"
```

```
print(sapply(transpose_m, class))
```

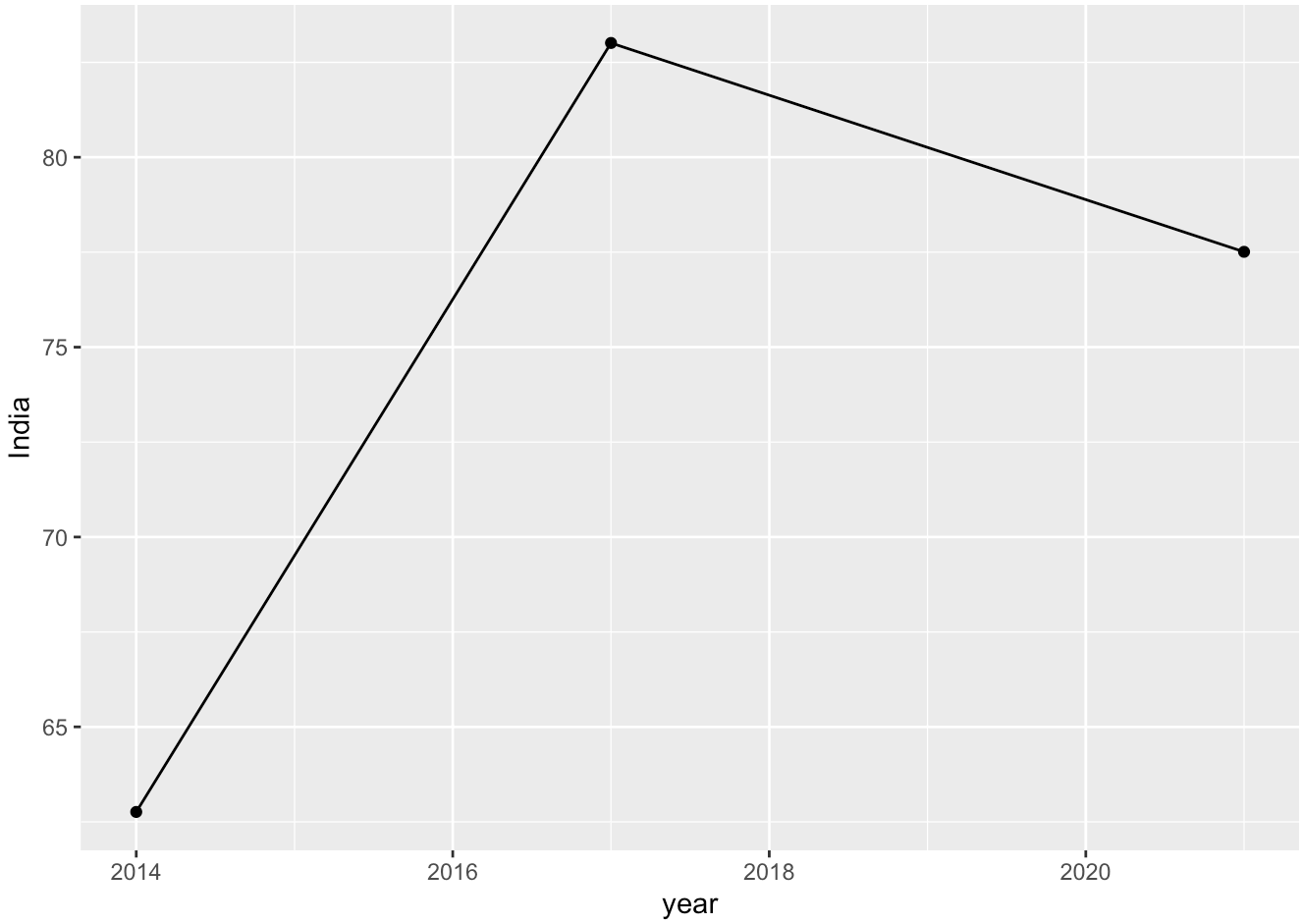
```
##        Brazil         India United States
##   "character"   "character"   "character"
```

```
#converts characters to numerical values
transpose_f$Brazil = as.numeric(transpose_f$Brazil)
transpose_f$India = as.numeric(transpose_f$India)
transpose_f$"United States" = as.numeric(transpose_f$"United States")
transpose_m$Brazil = as.numeric(transpose_m$Brazil)
transpose_m$India = as.numeric(transpose_m$India)
transpose_m$"United States" = as.numeric(transpose_m$"United States")

#creates a graph of female percentages
year <- c(2014, 2017,2021)
ggplot(data=transpose_f, aes(x=year, y=India, group=1)) +
  geom_line()+
  geom_point()
```



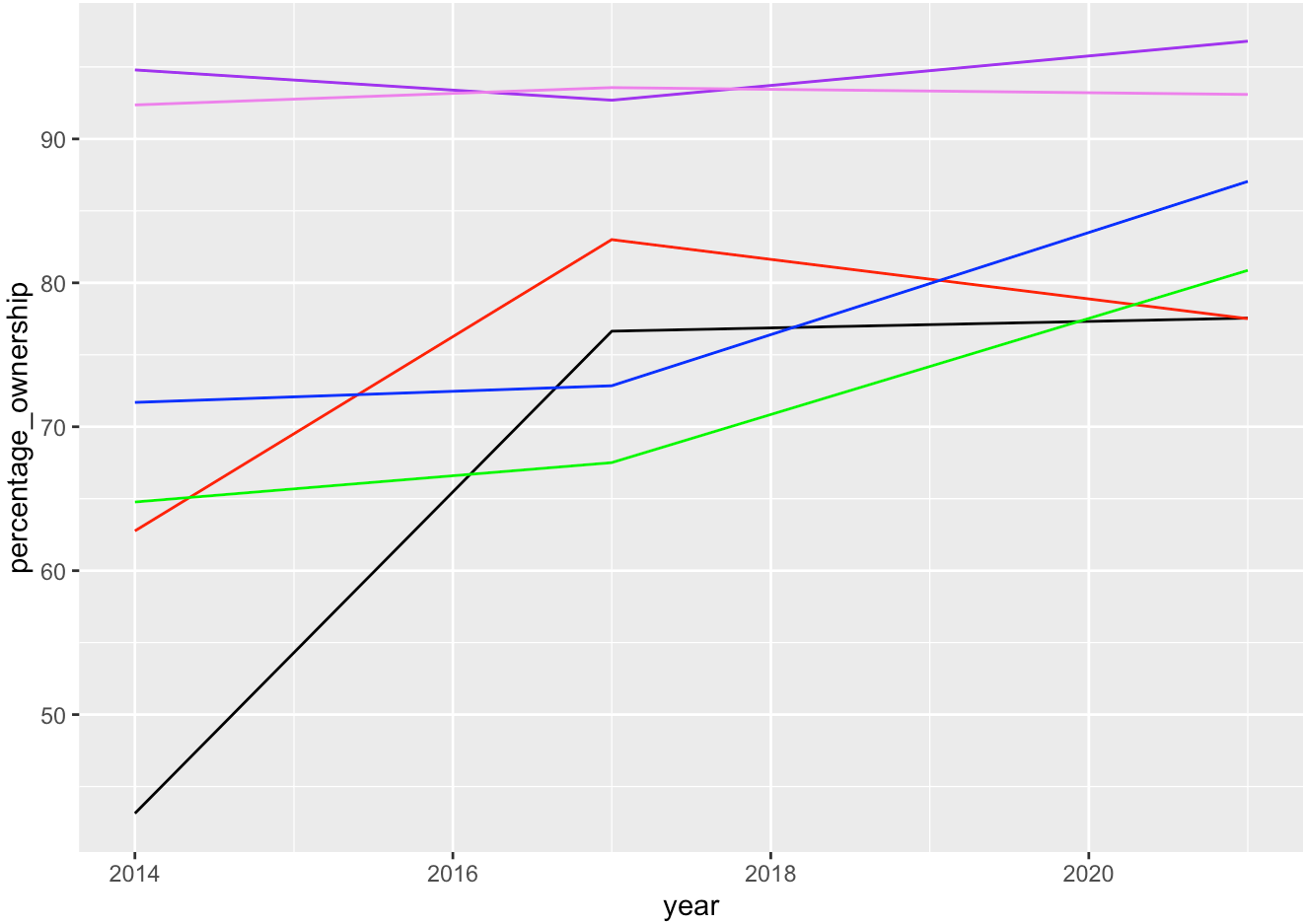

```
#creates a graph of male percentages
year <- c(2014, 2017,2021)
ggplot(data=transpose_m, aes(x=year, y=India, group=1)) +
  geom_line()+
  geom_point()
```


80
75
70
65
India
2014
2016
2018
2020
year

```
#renames titles to specify gender
transpose_f <- rename(transpose_f, "Brazil_f" = 1, "India_f" = 2, "United_States_f" = 3)
transpose_m <- rename(transpose_m, "Brazil_m" = 1, "India_m" = 2, "United_States_m" = 3)

#merges female and male data
transpose_m <- rownames_to_column(transpose_m, var="Year")
transpose_f <- rownames_to_column(transpose_f, var="Year")
acct_owner_by_gender <- merge(x = transpose_m, y = transpose_f, by = "Year", all.x = TRUE)
acct_owner_by_gender <- rename(acct_owner_by_gender, "United_States_f" = 7)

#creates plot of merged data
#1)
gfg_plot <- ggplot(acct_owner_by_gender, aes(x=year, y=percentage_ownership)) +
    geom_line(aes(y = India_f), color = "black") +
    geom_line(aes(y = India_m), color = "red") +
    geom_line(aes(y = Brazil_f), color = "green") +
    geom_line(aes(y = Brazil_m), color = "blue") +
    geom_line(aes(y = United_States_f), color = "purple") +
    geom_line(aes(y = United_States_m), color = "violet")
gfg_plot
```



#2) In the United States, account ownership at a financial institution or with a mobile-money-service provider is roughly equal for both males and females. However, in Brazil and India, males have a higher percentage of ownership. Between the years 2014 and 2020, percentage of both males and females with account ownership have increased in Brazil and India, while percentage of ownership in the United States has always been high. There is a change in slope for every demographic in the year 2017, indicating that something must have happened during that year that affected account ownership at a financial institution or with a mobile-money-service provider.

#3) Yes, even if you disregard a certain year, account percentages have still been increasing. When you look at the big picture, account percentages have overall increased from 2014 to 2020, so disregarding one year will not change the big picture. This is true for all demographics, regardless of country or gender. Except for the United States because in the United States, account percentage has always been very high. It is hard for such a high percentage to improve much more.

```
#load tidyverse
library(tidyverse)

setwd('~/Downloads/Account ownership at a financial institution or with a mobile-money-service provider, primary education or less (% of population ages 15+)')
primary <- read.csv("0e61522a-7eb7-4cd0-8428-b0ac66fca70f_Data.csv")
setwd('~/Downloads/Account ownership at a financial institution or with a mobile-money-service provider, secondary education or more (% of population ages 15+)')
secondary <- read.csv("7fed2679-8180-4c33-90cf-4b7efac2900d_Data.csv")

# Determine that R does consider primary to be data frame
is.data.frame(primary)
```

```
## [1] TRUE
```

```
# Determine that R does consider secondary to be data frame
is.data.frame(secondary)
```

```
## [1] TRUE
```

```
library(tidyverse)

#filters specific countries
primary <- primary %>%
  filter(Country.Name == "Austria" | Country.Name == "Canada" | Country.Name == "Finland" | Country.Name == "Italy" | Country.Name == "Spain")
secondary <- secondary %>%
  filter(Country.Name == "Austria" | Country.Name == "Canada" | Country.Name == "Finland" | Country.Name == "Italy" | Country.Name == "Spain")

rel_col <- which(colnames(primary)=="Country.Name" | colnames(primary)== "X2014..YR2014." | colnames(primary)== "X2017..YR2017." | colnames(primary) == "X2021..YR2021." )
rel_col
```

```
## [1]  3  9 12 16
```

```
#filters specific years
primary <- primary[rel_col] %>% rename("2014" = 2, "2017" = 3, "2021" = 4)
secondary <- secondary[rel_col] %>% rename("2014" = 2, "2017" = 3, "2021" = 4)

#swaps axes
transpose_p <- data.frame(t(primary[-1]))
colnames(transpose_p) <- primary[, 1]
transpose_s <- data.frame(t(secondary[-1]))
colnames(transpose_s) <- secondary[, 1]

#prints data type
print(sapply(transpose_p, class))
```

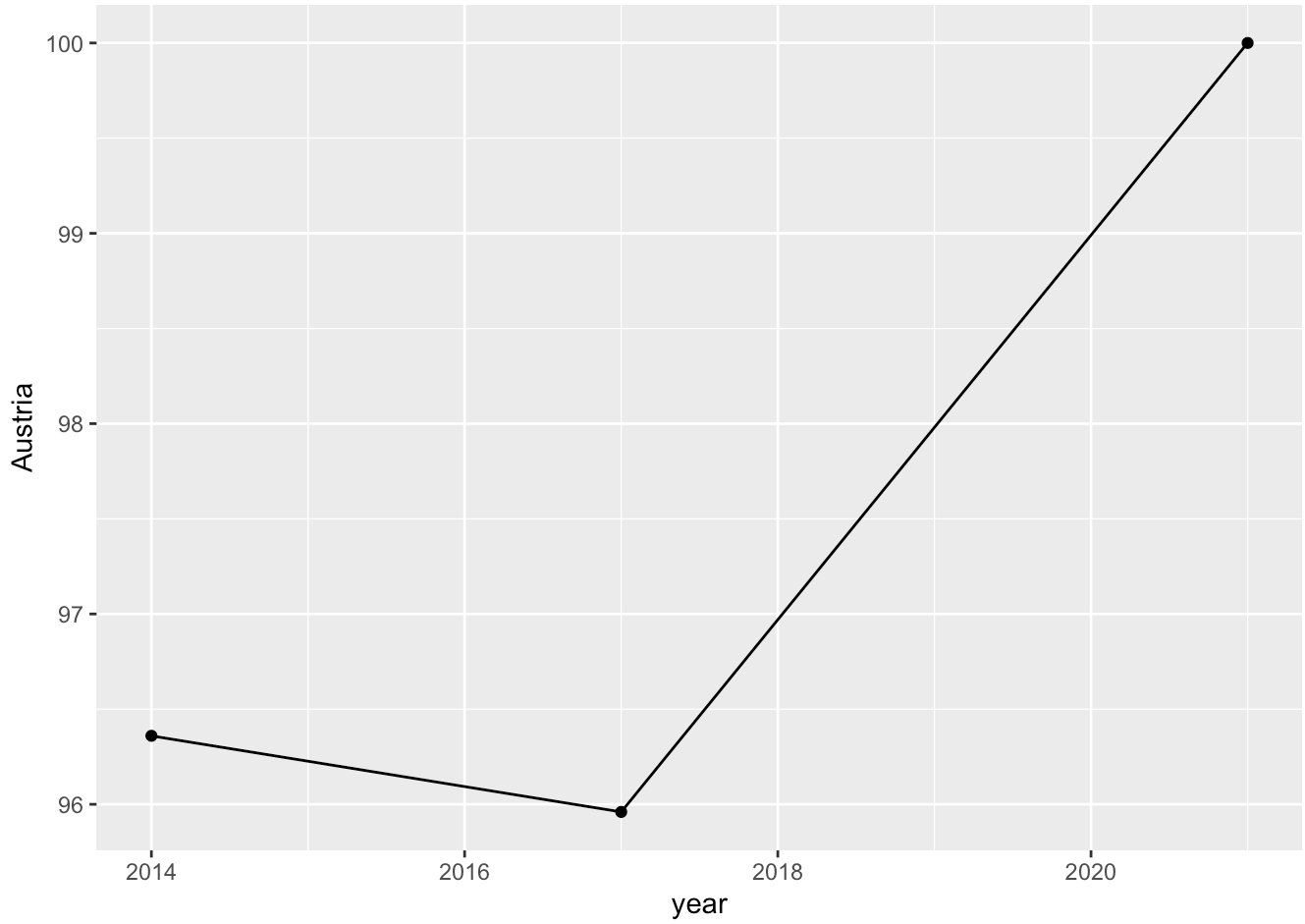
```
##     Austria      Canada     Finland       Italy       Spain
## "character" "character" "character" "character" "character"
```

```
print(sapply(transpose_s, class))
```

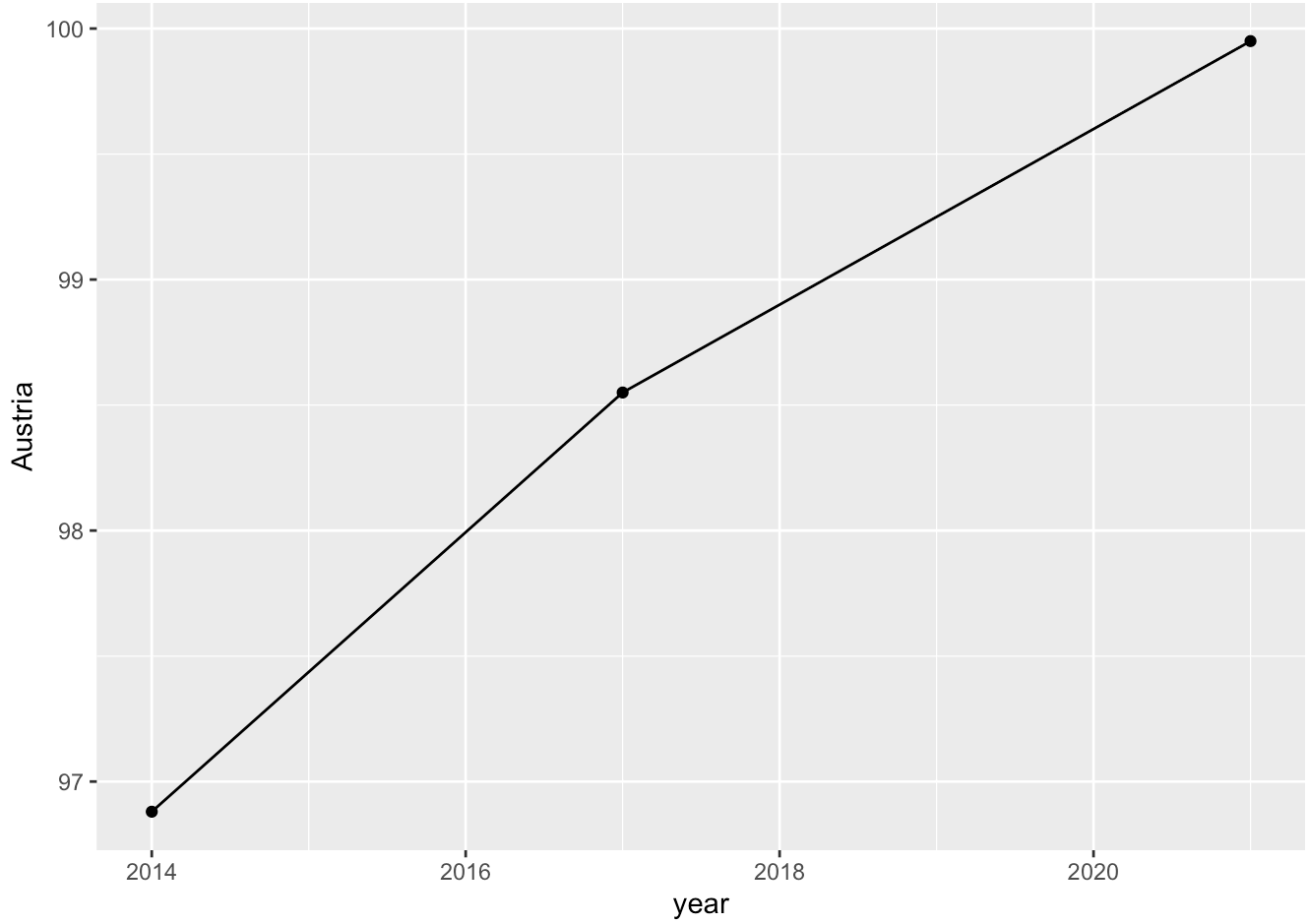
```
##     Austria      Canada     Finland       Italy       Spain
## "character" "character" "character" "character" "character"
```

```
#converts characters to numerical values
transpose_p$Austria = as.numeric(transpose_p$Austria)
transpose_p$Canada = as.numeric(transpose_p$Canada)
transpose_p$Finland = as.numeric(transpose_p$Finland)
transpose_p$Italy = as.numeric(transpose_p$Italy)
transpose_p$Spain = as.numeric(transpose_p$Spain)
transpose_s$Austria = as.numeric(transpose_s$Austria)
transpose_s$Canada = as.numeric(transpose_s$Canada)
transpose_s$Finland = as.numeric(transpose_s$Finland)
transpose_s$Italy = as.numeric(transpose_s$Italy)
transpose_s$Spain = as.numeric(transpose_s$Spain)

#creates a graph of primary percentages
year <- c(2014, 2017,2021)
ggplot(data=transpose_p, aes(x=year, y=Austria, group=1)) +
  geom_line()+
  geom_point()
```



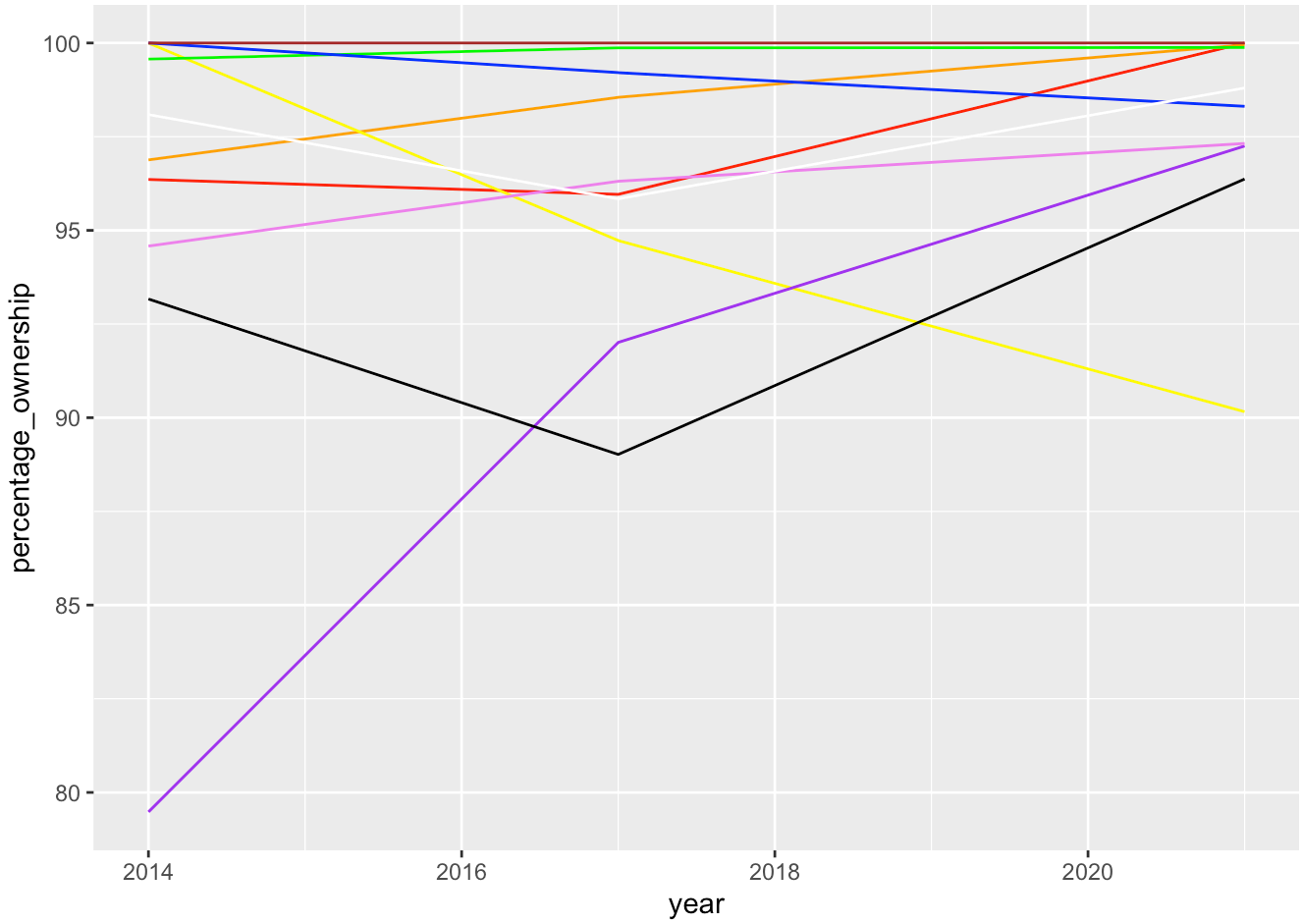

```
#creates a graph of secondary percentages
year <- c(2014, 2017,2021)
ggplot(data=transpose_s, aes(x=year, y=Austria, group=1)) +
  geom_line()+
  geom_point()
```


100
99
98
97
Austria
2014
2016
2018
2020
year

```
#renames titles to specify education level
transpose_p <- rename(transpose_p, "Austria_p" = 1, "Canada_p" = 2, "Finland_p" = 3, "Italy_p" = 4, "Spain_p" =
5)
transpose_s <- rename(transpose_s, "Austria_s" = 1, "Canada_s" = 2, "Finland_s" = 3, "Italy_s" = 4, "Spain_s" =
5)

#merges primary and secondary data
transpose_p <- rownames_to_column(transpose_p, var="Year")
transpose_s <- rownames_to_column(transpose_s, var="Year")
acct_owner_by_education <- merge(x = transpose_p, y = transpose_s, by = "Year", all.x = TRUE)
acct_owner_by_education <- rename(acct_owner_by_education)

#creates plot of merged data
#1)
gfg_plot <- ggplot(acct_owner_by_education, aes(x=year, y=percentage_ownership)) +
    geom_line(aes(y = Austria_p), color = "red") +
    geom_line(aes(y = Austria_s), color = "orange") +
    geom_line(aes(y = Canada_p), color = "yellow") +
    geom_line(aes(y = Canada_s), color = "green") +
    geom_line(aes(y = Finland_p), color = "blue") +
    geom_line(aes(y = Finland_s), color = "brown") +
    geom_line(aes(y = Italy_p), color = "purple") +
    geom_line(aes(y = Italy_s), color = "violet") +
    geom_line(aes(y = Spain_p), color = "black") +
    geom_line(aes(y = Spain_s), color = "white")
gfg_plot
```




```
#Conclusion) In all five countries, the percentage of those who attended secondary education had higher account o
wnership at a financial institution or with a mobile-money-service provider than those who only attended primary
education. Between 2014 and 2021, every demographic had an increase in account ownership, with the exception of t
hose who live in Canada in attended only primary education. By 2021, every demographic, regardless of nationality
or education, had a percentage ownership of at least 90%, which is very good. Those from Italy with primary educa
tion showed the most improvement, growing from ~80% account ownership to ~97% between 2014 and 2021. Overall, the
data shows a steady increase in financial inclusion in both primary and secondary education
```