The Farmer Dog

Amara Diallo

12/7/2020

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------------------------- tidyve
rse 1.3.0 --
```

```
## v ggplot2 3.3.2     v purrr   0.3.4
## v tibble  3.0.3     v dplyr   1.0.2
## v tidyr   1.1.2     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.5.0
```

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

```
library(ggplot2)
```

Looking at only the United States, France, and Spain, what was the median revenue in 2013? (Taking the three countries together in aggregate) *

```
`The.Farmer's.Dog` <- read.csv("C:/Users/User/Desktop/The Farmer's Dog.csv")

median_2013 <- `The.Farmer's.Dog` %>%
    select(country, year, revenue) %>%
    filter(country %in% c('US','FR', 'ES'), year =="2013") %>%

    #group_by(country, year) %>%
    summarize(median_3_country= median(revenue))

median_2013
```

```
##   median_3_country
## 1         93995.26
```

Was it more or less than 2012? *

```
median_2012 <- `The.Farmer's.Dog` %>%
    select(country, year, revenue) %>%
    filter(country %in% c('US','FR', 'ES'), year =="2012") %>%

    #group_by(country, year) %>%
    summarize(median_3_country= median(revenue))

median_2012
```

```
##   median_3_country
## 1         84380.38
```

What drove that difference? *

```
drove_2012 <- `The.Farmer's.Dog` %>%
    select(country, year, revenue,product_group) %>%
    filter(country %in% c('US','FR', 'ES'), year =="2012") %>%

    group_by(country,product_group) %>%
    summarize(revenue_m= median(revenue),
              sum_revenue = sum(revenue)) %>%
    mutate(product_group = product_group %>%
               as_factor() %>% fct_reorder(revenue_m)) %>%

    arrange(desc(revenue_m)) %>%
    ungroup()
```

```
## `summarise()` regrouping output by 'country' (override with `.groups` argument)
```

```
drove_2012
```

```
## # A tibble: 18 x 4
##    country product_group   revenue_m sum_revenue
##    <chr>   <fct>               <dbl>       <dbl>
##  1 FR      groceries         238991.     238991.
##  2 US      groceries         208650.     208650.
##  3 US      office supplies   207158.     207158.
##  4 ES      office supplies   206106.     206106.
##  5 US      automobile        201790.     201790.
##  6 FR      office supplies   201275.     201275.
##  7 ES      automobile        200856.     200856.
##  8 FR      automobile        190735.     190735.
##  9 ES      groceries         147368.     147368.
## 10 FR      furniture          21393.      21393.
## 11 FR      electronics        20353.      20353.
## 12 US      electronics        20175.      20175.
## 13 US      furniture          19864.      19864.
## 14 ES      furniture          19800.      19800.
## 15 ES      electronics        19505.      19505.
## 16 FR      clothing            1873.       1873.
## 17 US      clothing            1865.       1865.
## 18 ES      clothing            1082.       1082.
```

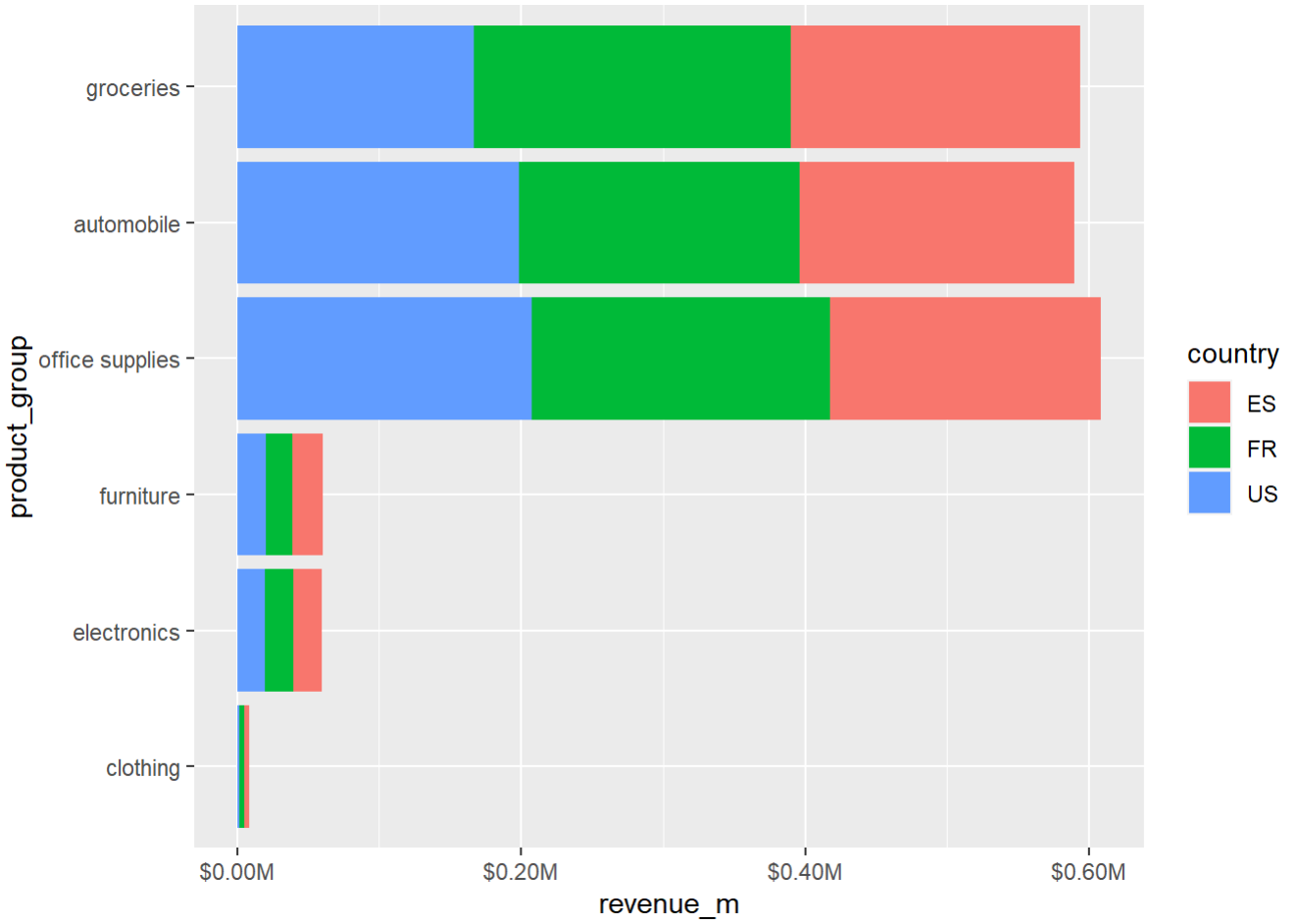
```
# drove_2013 <- `The.Farmer's.Dog` %>%
#     select(country, year, revenue, product_group) %>%
#     filter(country %in% c('US','FR', 'ES'), year =="2013") %>%
#
#     group_by( country,product_group) %>%
#     summarize(revenue_m= median(revenue),
#               sum_revenue = sum(revenue)) %>%
#     ungroup()
#
# drove_2013


drove_2013 <- `The.Farmer's.Dog` %>%
    select(country, year, revenue,product_group) %>%
    filter(country %in% c('US','FR', 'ES'), year =="2013") %>%

    group_by(country,product_group) %>%
    summarize(revenue_m= median(revenue),
              sum_revenue = sum(revenue)) %>%
    mutate(product_group = product_group %>%
               as_factor() %>% fct_reorder(revenue_m)) %>%

    arrange(desc(revenue_m)) %>%
    ungroup()
```

```
## `summarise()` regrouping output by 'country' (override with `.groups` argument)
```

```
drove_2013
```

```
## # A tibble: 18 x 4
##    country product_group   revenue_m sum_revenue
##    <chr>   <fct>               <dbl>       <dbl>
##  1 FR      groceries         223017.     223017.
##  2 FR      office supplies   209669.     209669.
##  3 US      office supplies   207829.     207829.
##  4 ES      groceries         203526.     203526.
##  5 US      automobile        198326.     198326.
##  6 FR      automobile        197660.     197660.
##  7 ES      automobile        193957.     193957.
##  8 ES      office supplies   191077.     191077.
##  9 US      groceries         167058.     167058.
## 10 ES      furniture          20933.      20933.
## 11 ES      electronics        20348.      20348.
## 12 US      furniture          20147.      20147.
## 13 US      electronics        19825.      19825.
## 14 FR      electronics        19615.      19615.
## 15 FR      furniture          19033.      19033.
## 16 FR      clothing            3346.       3346.
## 17 ES      clothing            3112.       3112.
## 18 US      clothing            1926.       1926.
```

```
drove_2013 %>%
    ggplot(aes(product_group, revenue_m, fill = country)) +
    geom_col() +
    scale_y_continuous(labels = scales::dollar_format(scale = 1e-6, suffix = "M")) +
    coord_flip()
```



Looking now at all countries, which product group had the biggest decline in median revenue from 2013 to 2014, by dollar amount? *

```
biggest_decline_prod <- `The.Farmer's.Dog` %>%
    select(country, year, revenue,product_group) %>%
    filter(year %in% c("2013", "2014")) %>%

    group_by(year, product_group) %>%
    summarize(revenue_m= median(revenue)) %>%

    arrange(desc(product_group)) %>%
    ungroup()
```

```
## `summarise()` regrouping output by 'year' (override with `.groups` argument)
```

```
biggest_decline_prod
```

```
## # A tibble: 12 x 3
##     year product_group   revenue_m
##    <int> <chr>                <dbl>
##  1  2013 office supplies   201386.
##  2  2014 office supplies   199426.
##  3  2013 groceries         210322.
##  4  2014 groceries         200448.
##  5  2013 furniture          20037.
##  6  2014 furniture          20284.
##  7  2013 electronics        20078.
##  8  2014 electronics        19631.
##  9  2013 clothing            2054.
## 10  2014 clothing            1634.
## 11  2013 automobile        197911.
## 12  2014 automobile        197353.
```

Looking now at all countries, which product group had the biggest decline in median revenue from 2013 to 2014, by percent change?

```
# biggest_decline_prod %>%
#     mutate(revenue_m_dif = lag(revenue_m, n =1)) %>%
#     mutate(revenue_m_dif = case_when(
#         is.na(revenue_m_dif) ~ revenue_m,
#         TRUE ~ revenue_m_dif
#     )) %>%
#
#     mutate(diff_1 = revenue_m - revenue_m_dif) %>%
#     mutate(pct_diff_1 = diff_1/revenue_m_dif) %>%
#     mutate(pct_diff_1_chr = scales::percent(pct_diff_1, accuracy = 3)) %>%
#
#     select(year, product_group,revenue_m, pct_diff_1_chr) %>%
#     filter(product_group=='clothing') %>%
#     arrange(desc(product_group))


biggest_decline_prod %>%
    group_by(product_group) %>%
    mutate(pct_change = (revenue_m/lag(revenue_m)-1) * 100) %>%
    mutate(pct_change = scales::percent(pct_change,scale = 3)) %>%

    filter(product_group=='clothing')
```

```
## # A tibble: 2 x 4
## # Groups:   product_group [1]
##    year product_group revenue_m pct_change
##   <int> <chr>             <dbl> <chr>
## 1  2013 clothing          2054. <NA>
## 2  2014 clothing          1634. -61%
```

```
    #arrange(year, .by_group = TRUE)
```