Case Study - Online Sales

Ahmet Selimhan ŞAHİN

20 11 2021

Homework 1

Downloading the required packages

```
## Messages and warnings are suspended to prevent reader to be disturbed
library(knitr)
library(ggplot2)
library(rworldmap)
library(dplyr)
```

```
load("F:/ders/3rd_class/IE421 Data Science for Engineers/week6/salesData.Rdata")
```

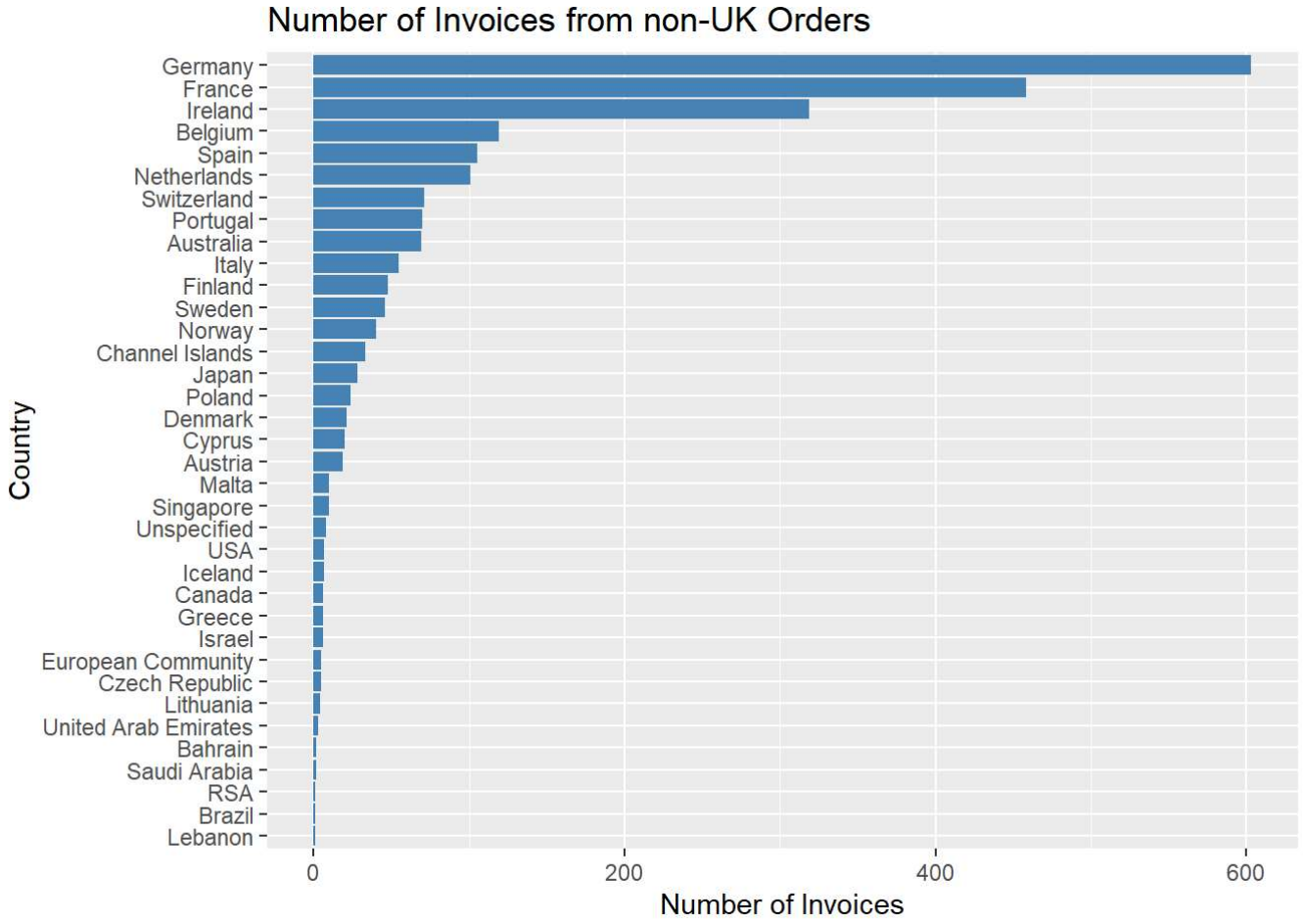
Question 1 Visualizing Number of Invoices

```
invoiceData = salesData %>%
  group_by(InvoiceNo , Country) %>%
  summarize(
    count = n()
  )

invoiceData
```

```
## # A tibble: 22,190 x 3
## # Groups:   InvoiceNo [22,190]
##    InvoiceNo Country        count
##    <fct>     <fct>          <int>
##  1 536365    United Kingdom     7
##  2 536366    United Kingdom     2
##  3 536367    United Kingdom    12
##  4 536368    United Kingdom     4
##  5 536369    United Kingdom     1
##  6 536370    France            20
##  7 536371    United Kingdom     1
##  8 536372    United Kingdom     2
##  9 536373    United Kingdom    16
## 10 536374    United Kingdom     1
## # ... with 22,180 more rows
```

```
invoiceData %>%
  filter(Country != "United Kingdom") %>%
  ggplot() + geom_bar(mapping = aes(reorder(Country , table(Country)[Country])) , fill = "ste
elblue") +
  labs(x = "Country" , y = "Number of Invoices" , title = "Number of Invoices from non-UK Ord
ers") + coord_flip()
```



Question 2 Total Revenue from Countries

```
TotalPrice = mutate(salesData , TotalPrice =
                      Quantity * UnitPrice)
head(TotalPrice , 15)
```

```
##    InvoiceNo StockCode                         Description Quantity
## 1     536365    85123A  WHITE HANGING HEART T-LIGHT HOLDER        6
## 2     536365     71053                 WHITE METAL LANTERN        6
## 3     536365    84406B      CREAM CUPID HEARTS COAT HANGER        8
## 4     536365    84029G KNITTED UNION FLAG HOT WATER BOTTLE        6
## 5     536365    84029E      RED WOOLLY HOTTIE WHITE HEART.        6
## 6     536365     22752        SET 7 BABUSHKA NESTING BOXES        2
## 7     536365     21730   GLASS STAR FROSTED T-LIGHT HOLDER        6
## 8     536366     22633              HAND WARMER UNION JACK        6
## 9     536366     22632           HAND WARMER RED POLKA DOT        6
## 10    536367     84879       ASSORTED COLOUR BIRD ORNAMENT       32
## 11    536367     22745          POPPY'S PLAYHOUSE BEDROOM         6
## 12    536367     22748           POPPY'S PLAYHOUSE KITCHEN        6
## 13    536367     22749   FELTCRAFT PRINCESS CHARLOTTE DOLL        8
## 14    536367     22310             IVORY KNITTED MUG COSY         6
## 15    536367     84969  BOX OF 6 ASSORTED COLOUR TEASPOONS        6
##            InvoiceDate UnitPrice CustomerID        Country TotalPrice
## 1  2010-12-01 08:26:00      2.55      17850 United Kingdom      15.30
## 2  2010-12-01 08:26:00      3.39      17850 United Kingdom      20.34
## 3  2010-12-01 08:26:00      2.75      17850 United Kingdom      22.00
## 4  2010-12-01 08:26:00      3.39      17850 United Kingdom      20.34
## 5  2010-12-01 08:26:00      3.39      17850 United Kingdom      20.34
## 6  2010-12-01 08:26:00      7.65      17850 United Kingdom      15.30
## 7  2010-12-01 08:26:00      4.25      17850 United Kingdom      25.50
## 8  2010-12-01 08:28:00      1.85      17850 United Kingdom      11.10
## 9  2010-12-01 08:28:00      1.85      17850 United Kingdom      11.10
## 10 2010-12-01 08:34:00      1.69      13047 United Kingdom      54.08
## 11 2010-12-01 08:34:00      2.10      13047 United Kingdom      12.60
## 12 2010-12-01 08:34:00      2.10      13047 United Kingdom      12.60
## 13 2010-12-01 08:34:00      3.75      13047 United Kingdom      30.00
## 14 2010-12-01 08:34:00      1.65      13047 United Kingdom       9.90
## 15 2010-12-01 08:34:00      4.25      13047 United Kingdom      25.50
```

```
top10_countries = TotalPrice %>%
  group_by(Country) %>%
  summarize(
    countrywise_totalrevenue =         sum(TotalPrice) ) %>%
  arrange(desc(countrywise_totalrevenue)) %>%
  top_n(n = 10 , wt = countrywise_totalrevenue)

top10_countries
```

```
## # A tibble: 10 x 2
##    Country        countrywise_totalrevenue
##    <fct>                             <dbl>
##  1 United Kingdom                6767873.
##  2 Netherlands                    284662.
##  3 Ireland                        250285.
##  4 Germany                        221698.
##  5 France                         196713.
##  6 Australia                      137077.
##  7 Switzerland                     55739.
##  8 Spain                           54775.
##  9 Belgium                         40911.
## 10 Sweden                          36596.
```

```
kable(top10_countries , align = c("c" , "c") , col.names = c("Country" , "Total Revenue") , c
aption = "Top 10"  )
```

Top 10

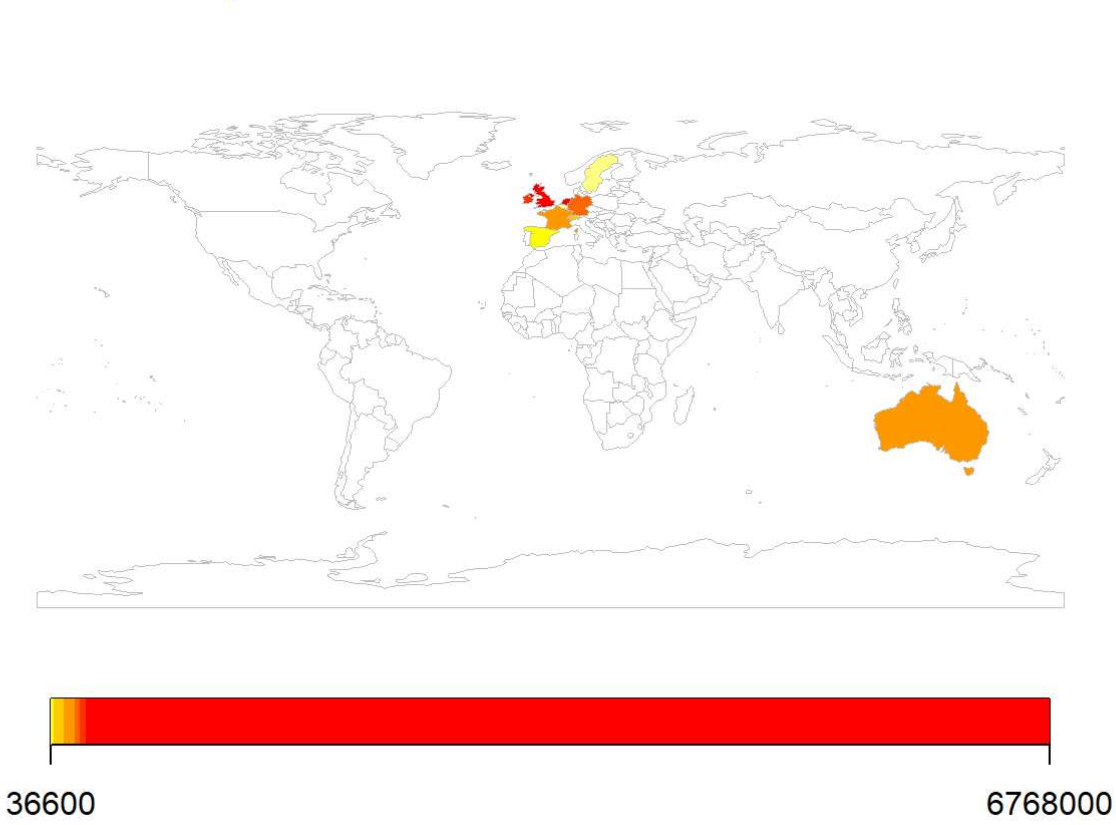
Country	Total Revenue
United Kingdom	6767873.39
Netherlands	284661.54
Ireland	250285.22
Germany	221698.21
France	196712.84
Australia	137077.27
Switzerland	55739.40
Spain	54774.58
Belgium	40910.96
Sweden	36595.91

Top 10 Countries on the Map

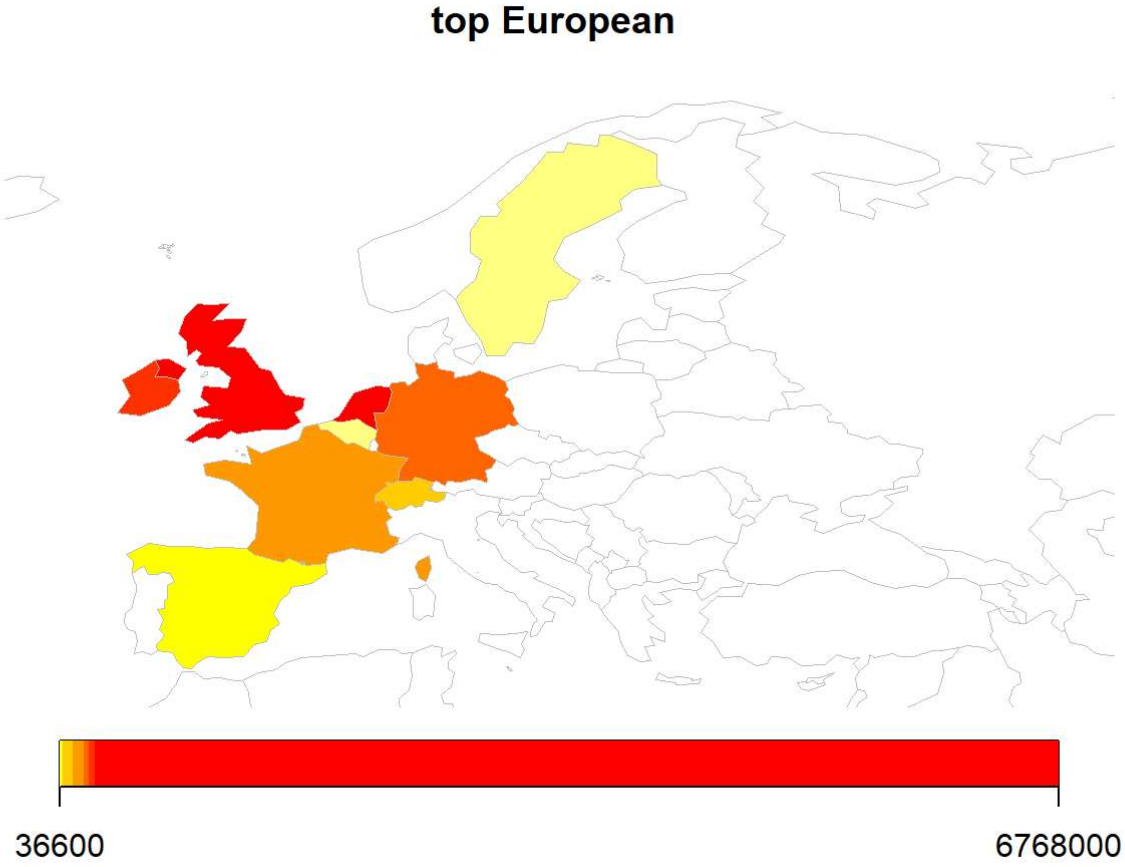
```
n <- joinCountryData2Map(top10_countries , joinCode = "NAME" , nameJoinColumn = "Country")
```

```
## 10 codes from your data successfully matched countries in the map
## 0 codes from your data failed to match with a country code in the map
## 233 codes from the map weren't represented in your data
```

```
mapCountryData(n , nameColumnToPlot = "countrywise_totalrevenue" , mapTitle = "Top 10 Countri
es Based on total revenue")
```




```
mapCountryData(n , nameColumnToPlot = "countrywise_totalrevenue" , mapRegion = "Europe" , map
Title = "top European ")
```



Question 3 Most Valued Products and Customers

```
TotalPrice %>%
  group_by(Description) %>%
  summarize(
    count = n()
  )
```

```
## # A tibble: 3,896 x 2
##    Description                         count
##    <fct>                               <int>
##  1 WHITE HANGING HEART T-LIGHT HOLDER   2070
##  2 WHITE METAL LANTERN                   259
##  3 CREAM CUPID HEARTS COAT HANGER        250
##  4 KNITTED UNION FLAG HOT WATER BOTTLE   330
##  5 RED WOOLLY HOTTIE WHITE HEART.        333
##  6 SET 7 BABUSHKA NESTING BOXES          306
##  7 GLASS STAR FROSTED T-LIGHT HOLDER     129
##  8 HAND WARMER UNION JACK                477
##  9 HAND WARMER RED POLKA DOT              18
## 10 ASSORTED COLOUR BIRD ORNAMENT        1418
## # ... with 3,886 more rows
```

```
data_named = TotalPrice %>%
  group_by(Description) %>%
  summarize(
    value_product = sum(TotalPrice)) %>%
  arrange(desc(value_product)) %>%
  top_n(n = 20, wt = value_product)

kable(data_named , align = c("c" , "c") , col.names = c("Product" , "Value") , caption = "Top
20")
```

Top 20

Product	Value
REGENCY CAKESTAND 3 TIER	132870.40
WHITE HANGING HEART T-LIGHT HOLDER	93823.85
JUMBO BAG RED RETROSPOT	83236.76
PARTY BUNTING	67687.53
POSTAGE	66710.24
ASSORTED COLOUR BIRD ORNAMENT	56499.22
RABBIT NIGHT LIGHT	51137.80
CHILLI LIGHTS	45936.81
PAPER CHAIN KIT 50’S CHRISTMAS	41500.48

Product	Value
PICNIC BASKET WICKER 60 PIECES	39619.50
BLACK RECORD COVER FRAME	39009.38
JUMBO BAG PINK POLKADOT	36473.01
SPOTTY BUNTING	35056.44
DOORMAT KEEP CALM AND COME IN	34312.60
WOOD BLACK BOARD ANT WHITE FINISH	34307.06
SET OF 3 CAKE TINS PANTRY DESIGN	32607.80
JAM MAKING SET WITH JARS	31658.47
JUMBO BAG STRAWBERRY	30195.71
HEART OF WICKER LARGE	28331.45
VICTORIAN GLASS HANGING T-LIGHT	28152.86

Number of Customers

```
TotalPrice %>%
  group_by(CustomerID) %>%
  summarize(
    count = n()
  )
```

```
## # A tibble: 4,372 x 2
##    CustomerID count
##         <int> <int>
##  1      12346     2
##  2      12347   182
##  3      12348    31
##  4      12349    73
##  5      12350    17
##  6      12352    95
##  7      12353     4
##  8      12354    58
##  9      12355    13
## 10      12356    59
## # ... with 4,362 more rows
```

```
top20_cust = TotalPrice %>%
  group_by(CustomerID) %>%
  summarize(customer_amount = sum(TotalPrice)) %>%
  arrange(desc(customer_amount)) %>%
  top_n(n = 20 , wt = customer_amount)


kable(top20_cust , align = c("c" , "c"))
```

CustomerID	customer_amount
14646	279489.02
18102	256438.49
17450	187482.17
14911	132572.62
12415	123725.45
14156	113384.14
17511	88125.38
16684	65892.08
13694	62653.10
15311	59419.34
13089	57385.88
14096	57120.91
15061	54228.74
17949	52750.84
15769	51823.72
16029	50992.61
14298	50862.44
14088	50415.49
17841	40340.78
13798	36351.42