Internet sports betting in 2005
Datamart creation and insights

Abdon Ahuile,Annatol Bannem, Charles Payet

Abstract

For this project we used the different tables provided from an internet sports betting website in order to create a final data mart which summarizes the behaviour and also includes the demographic behaviour of the clients. The final output isa table of one row per User Id with 51 different variables.

Libraries

The required libraries for making these project are the following:

```
library(haven)
```

```
## Warning: package 'haven' was built under R version 3.6.2
```

```
library(dplyr)
library(tidyr)
library(data.table)
library(ggplot2)
library(shiny)
```

```
## Warning: package 'shiny' was built under R version 3.6.2
```

Analysis

Data Preprocessing

The tables used to make the final merge in order to make the final datamart are "Demograpics", "Analytic-DataInternetGambling" which where provided beforehand? This where merged with tables "pokerchipfinal" and "dailyfinal" which come from "RawDataIIIPokerChipConversions" and "RawDataIIUserDailyAggregation" which contained repeated User Ids, therefore they where summarized to have one observation per client.

Relevant analysis

As it can be observed the German speaking users have accumulated the most amount of bets during the period of the data provided.

In the graph above it can be noticed that in every country in which a betting service provider is being used, this service is predominatly used by males (at least for the users which have given their gender information).
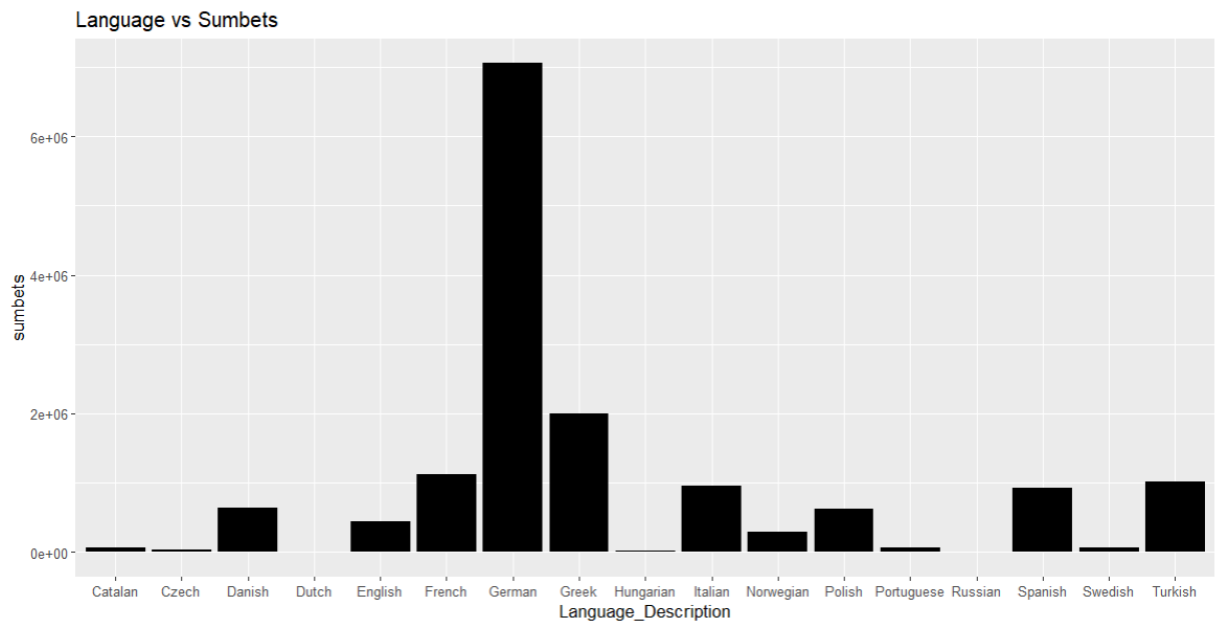
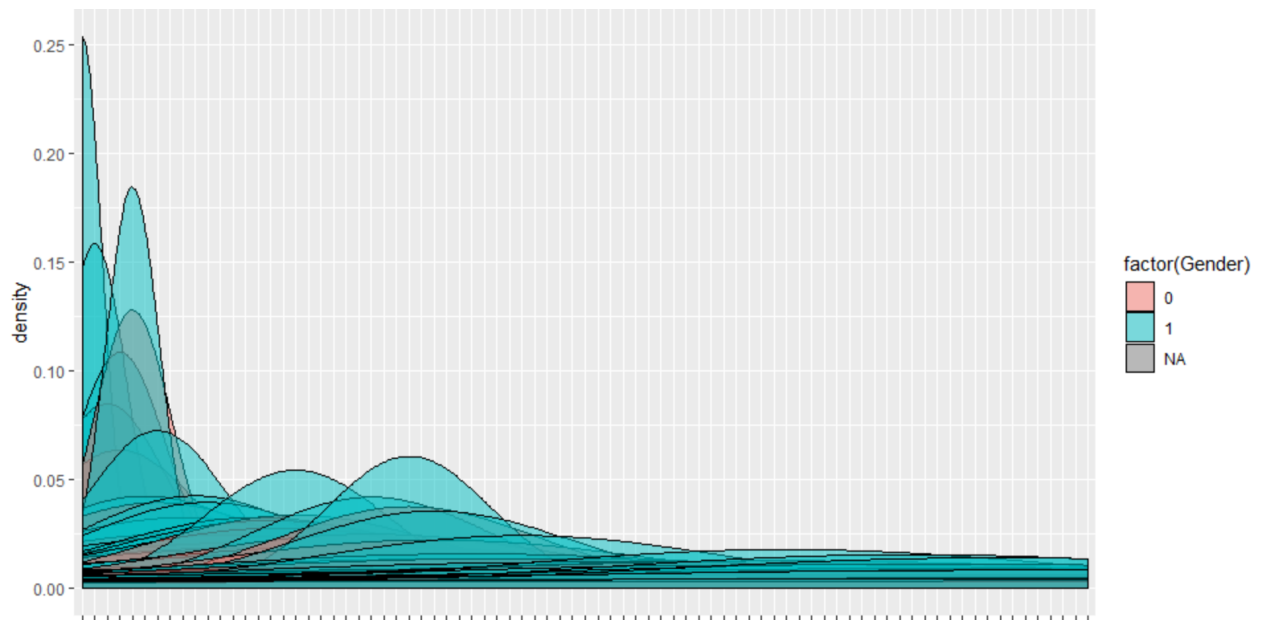Figure 1: Total bets grouped by users who speak different languages .

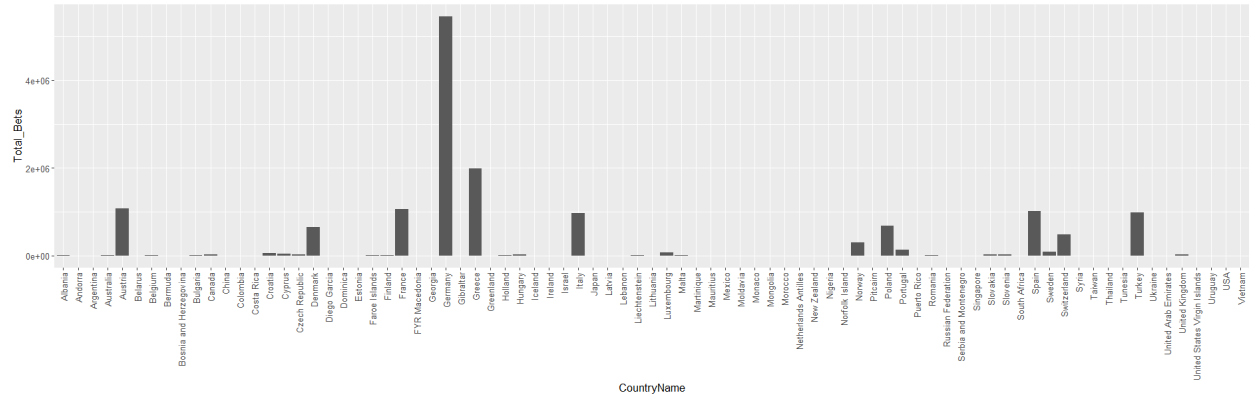Figure 2: Density graph of Gender and country.

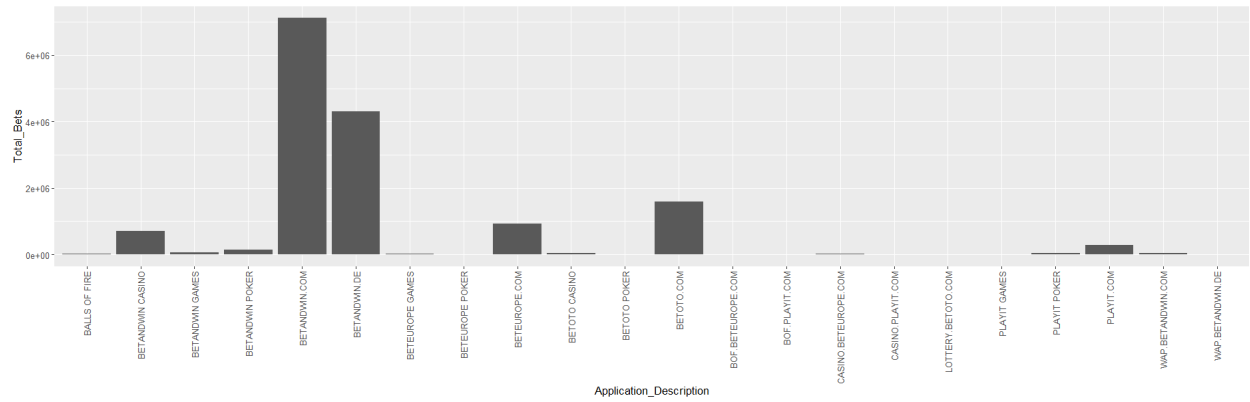Figure 3: Bets by Country.

Figure 4: Bets by Applications.

With the above graph, we can notice that Germany has the highest betting habit among all the countries.

With the help the graph, we can easily say that the most common app used by Gamblers is BETAND-WIN.COM followed by BETANDWIN.DE
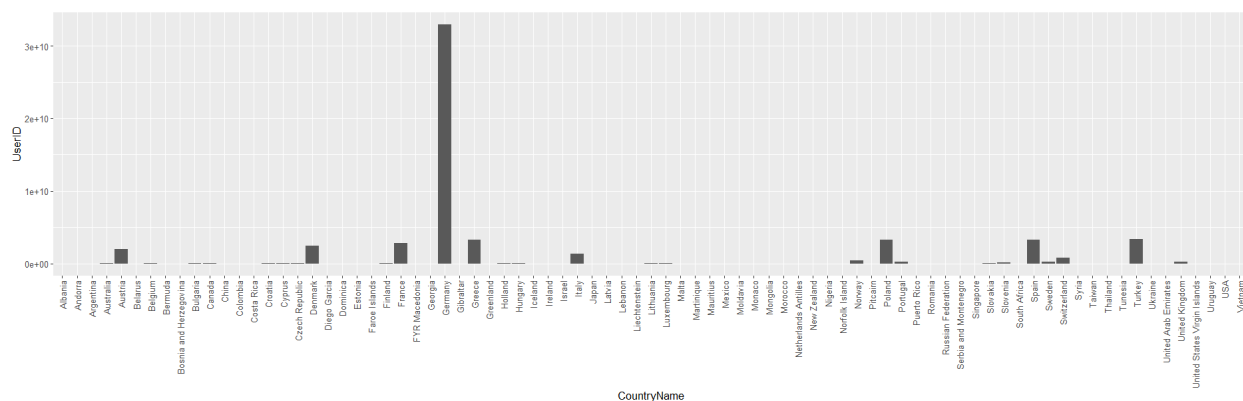

Figure 5: Avg of Customers by Country

On the graph above, the betting company can notice that the majority of their customers are coming from Germany.
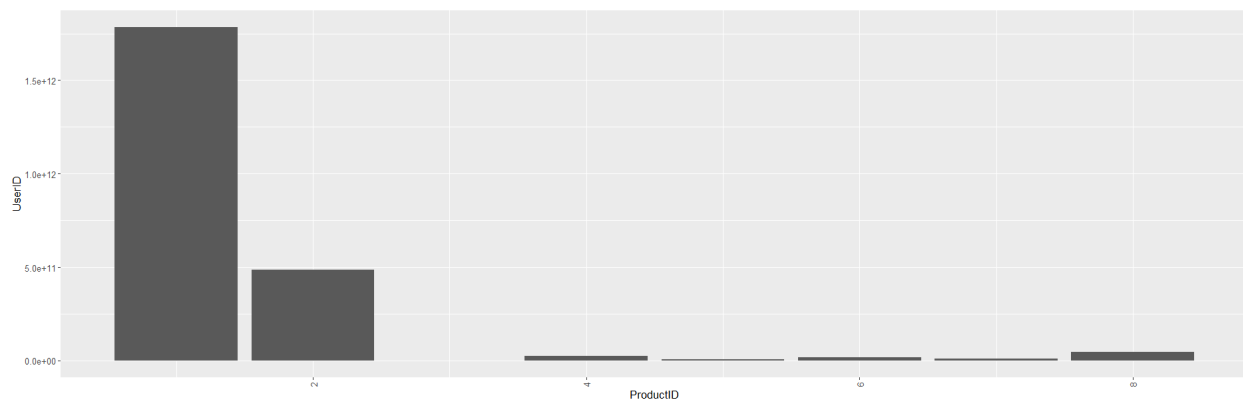

Figure 6: Products consumed by Customers

On this graph, the betting company can launch a marketing process on the product 1(Sports book fixed-odd) since it's the Customers favorite.

On this graph, the betting comapny can try to target the customers around 35 years old and below (With a peak at 31 years old) since they are the most interrested in the betting activity.

On this graph, we can notice that Denmark is the country with the highest Pokerchip Buying behaviour. The betting company can use this to set their future selling target.
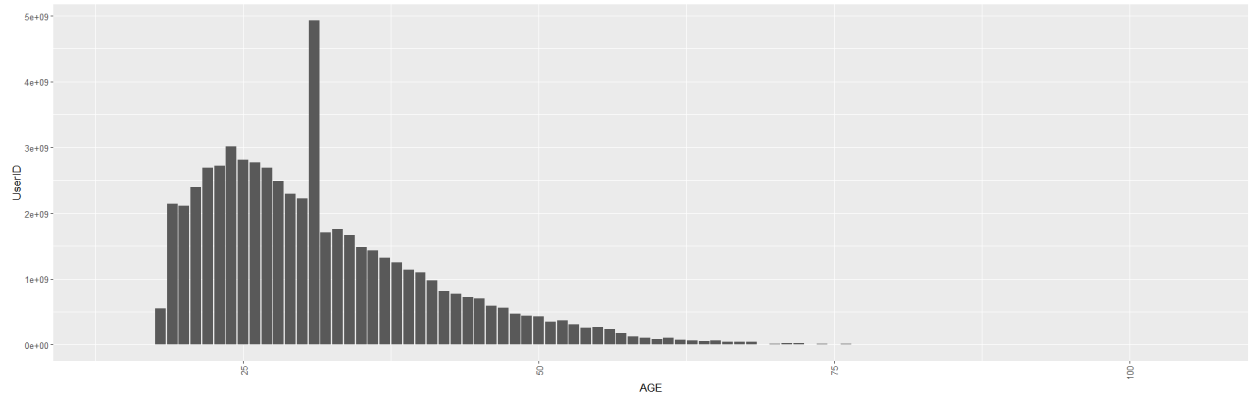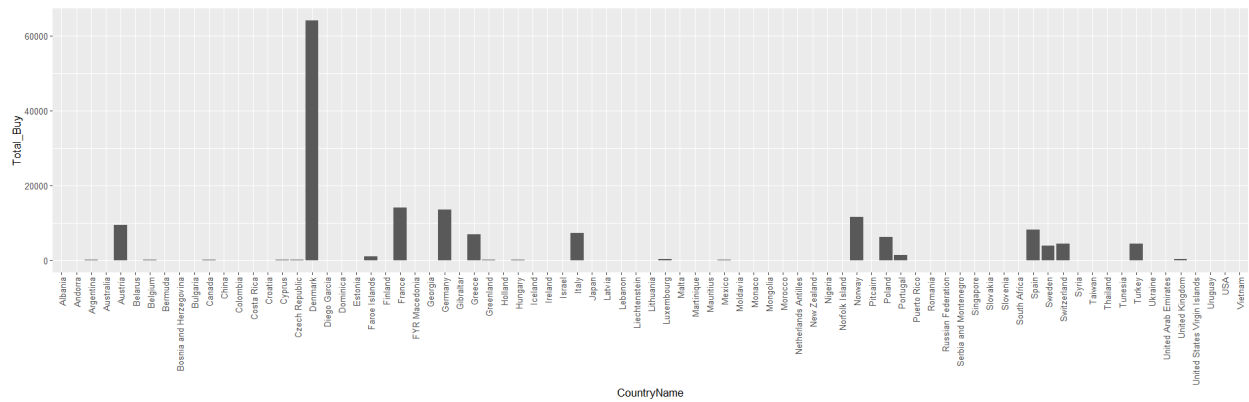
Figure 7: Customers Age

Figure 8: What country is buying pokerchip