

# Internet Sports Gambling Activity Analysis & Marketing Data Mart

Rohit Bhalerao, Marie Reine, Alessio Rosato

December 23, 2020

## Project Description

The data from the betting company BWin includes demographic data of customers, transaction behavior and wins, stakes and bets for different products (games). Based on this data, a datamart is created with one record per customer. This helps us analyze the behaviour of clients which allows the company to pitch marketing campaigns accordingly. Based on the datamart, marketing metrics are calculated which are used to show insights on the shiny application.

## 1. Data Description

The following data was provided by BWin:

- **User Demographic** - Contains user data like Gender, Country, Language, Application used and the first date of activity for various activity types.
- **UserDailyAggregation** - Contains the bets, stakes and amount winnings for each user for each product type.
- **PokerChipConversions** - Contains the transaction of amount sold or bought for each user.
- **Analytic Dataset** - Contains actual user behaviour data including the first and last activity dates for different types of products.

## 2. Data Manipulation and Creating Datamart

For the datamart, the following changes were made to the original data :

- The dates which were originally characters were changed to date values.
- The pokerchip data was summarized to only have the first and last date of transaction.
- New columns are created for each type of product in the pokerchip data.
- In user aggregation data, we place all negative values by 0
- A wider pivot table is created from User Aggregation table with columns for each variable.
- We merge the pivot tables that are created in the dataframe.
- We replace all NULL and NA values by 0, and also remove columns with many 0's. One row with no Gender value is dropped.
- In the merged data, replace all integers with the actual value. The values are in the data documentation. Some columns are dropped because they are not required or are redundant.

The final datamart has the following variables:

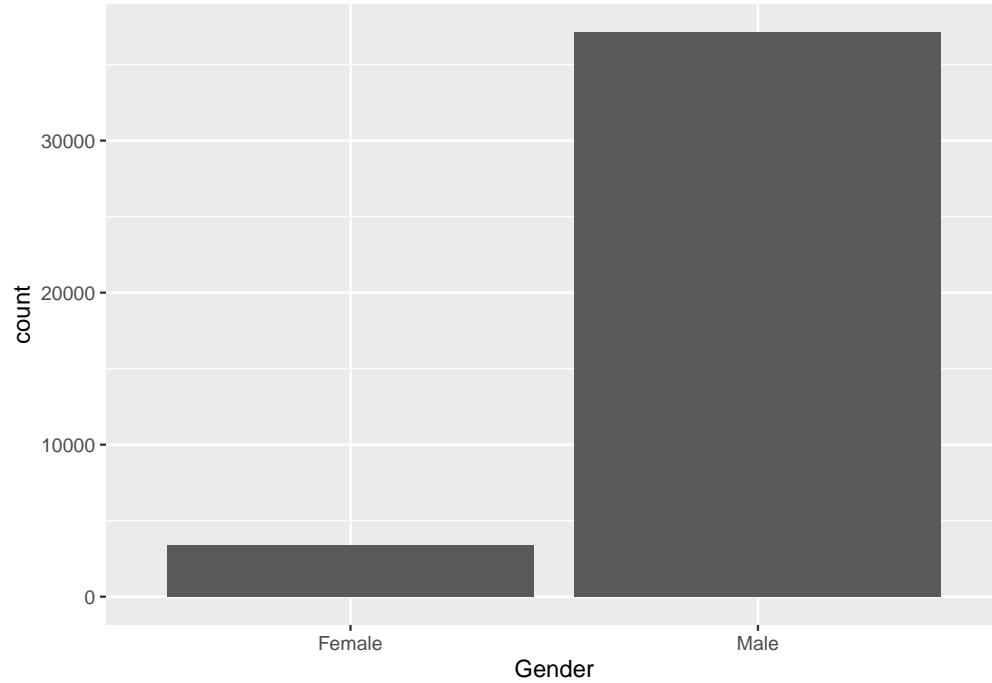
```
## [1] "X"
## [2] "UserID"
## [3] "Gender"
## [4] "Fixed.Odds.Stake..Total."
## [5] "Fixed.Odds.Winnings..Total."
## [6] "Fixed.Odds.Bets..Total."
## [7] "Live.Action.Stake..Total."
## [8] "Live.Action.Winnings..Total."
## [9] "Live.Action.Bets..Total."
## [10] "BettingApplication"
## [11] "Number.of.Buying.Transactions"
## [12] "Number.of.Selling.Transactions"
## [13] "Total.Amount..Buying."
## [14] "Total.Amount..Selling."
## [15] "Sports.Book.Fixed.Odd..Count."
## [16] "Sports.Book.Live.Action..Count."
## [17] "Casino.Chartwell..Count."
## [18] "Games.VS..Count."
## [19] "Games.BWIN..Count."
## [20] "Casino.BossMedia..Count."
## [21] "Supertoto..Count."
## [22] "Sports.Book.Fixed.Amount..Total.Stakes."
## [23] "Sports.Book.Live.Action..Total.Stakes."
## [24] "Casino.Chartwell..Total.Stakes."
## [25] "Games.VS..Total.Stakes."
## [26] "Games.BWIN..Total.Stakes."
## [27] "Casino.BossMedia..Total.Stakes."
## [28] "Supertoto..Total.Stakes."
## [29] "Sports.Book.Fixed.Odds..Total.Winnings."
## [30] "Sports.Book.Live.Action..Total.Winnings."
## [31] "Casino.Chartwell..Total.Winnings."
## [32] "Games.VS..Total.Winnings."
## [33] "Games.BWIN..Total.Winnings."
## [34] "Casino.BossMedia..Total.Winnings."
## [35] "Supertoto..Total.Winnings."
## [36] "Sports.Book.Fixed.Odds..Total.Bets."
## [37] "Sports.Book.Live.Action..Total.Bets."
## [38] "Casino.Chartwell..Total.Bets."
## [39] "Games.VS..Total.Bets."
## [40] "Games.BWIN..Total.Bets."
## [41] "Casino.BossMedia..Total.Bets."
## [42] "Supertoto..Total.Bets."
## [43] "Country"
## [44] "Language"
## [45] "TimeToDeposit"
```

### 3. Basic Data Analysis

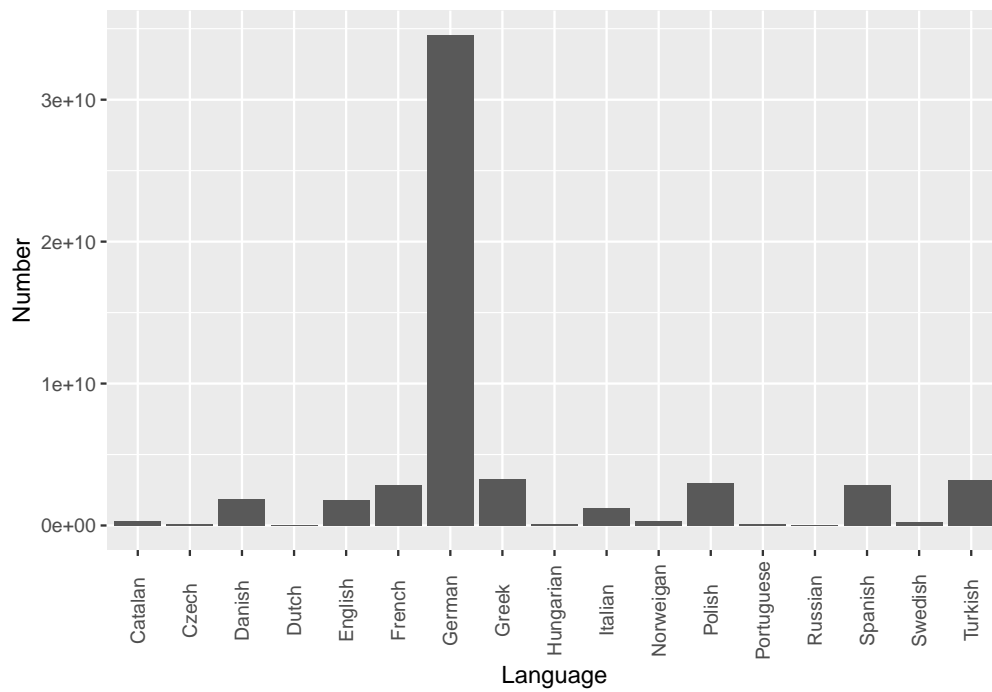
- 40499 Users in the datamart
- Higher proportion of Males compared to Females

- Betandwin.Com has highest number of users
- Denmark, France and Germany have highest number of poker players

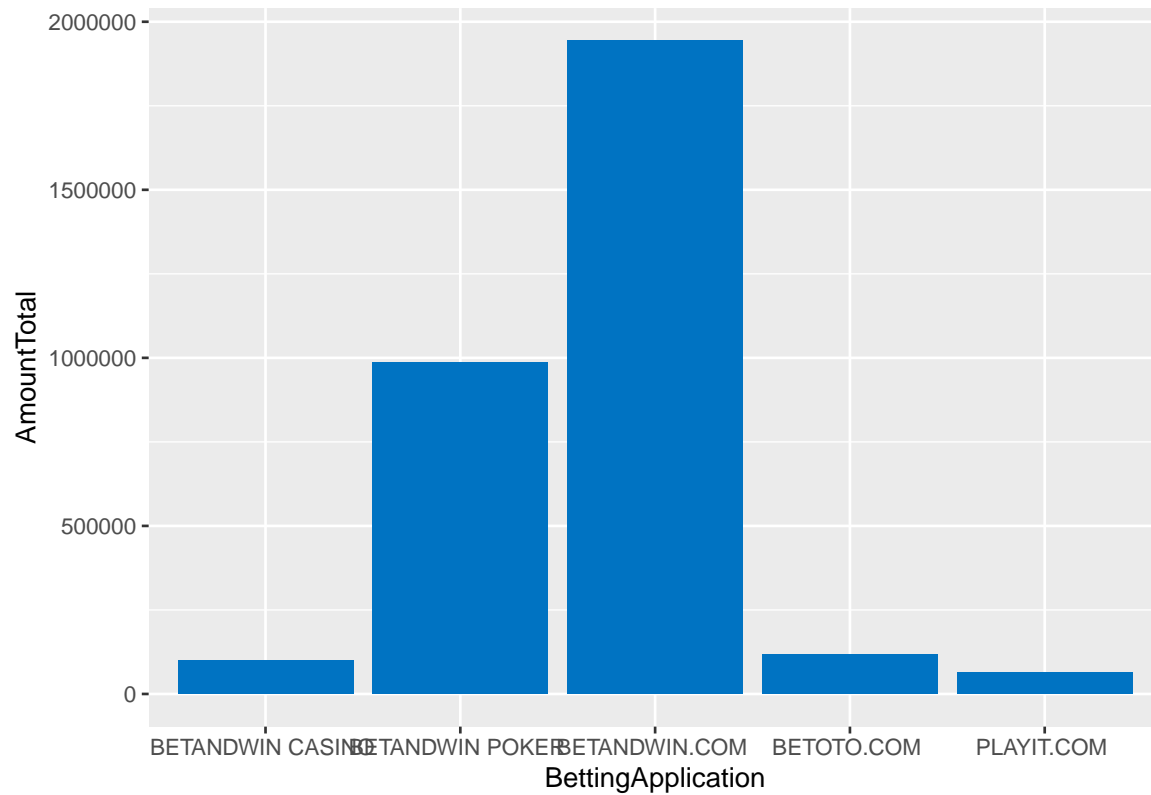
**Gender** 92% of gamblers are men.



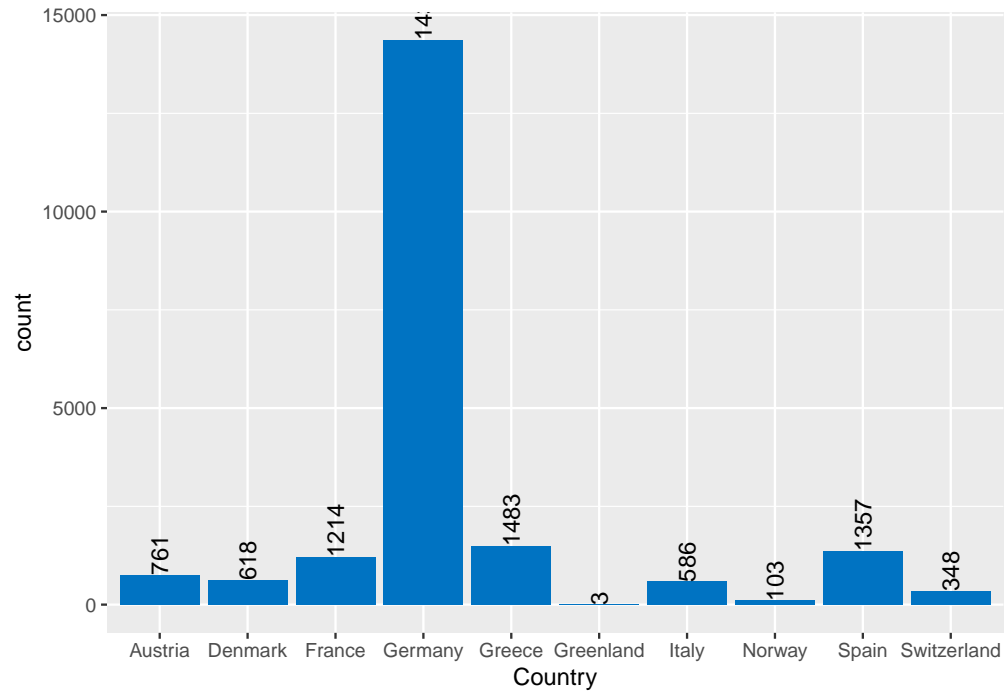
**Language** Most gamblers are German speakers.



**Betting Applications** Betandwin.Com is the most common application.



**Countries** Most of the poker players are from Germany.



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

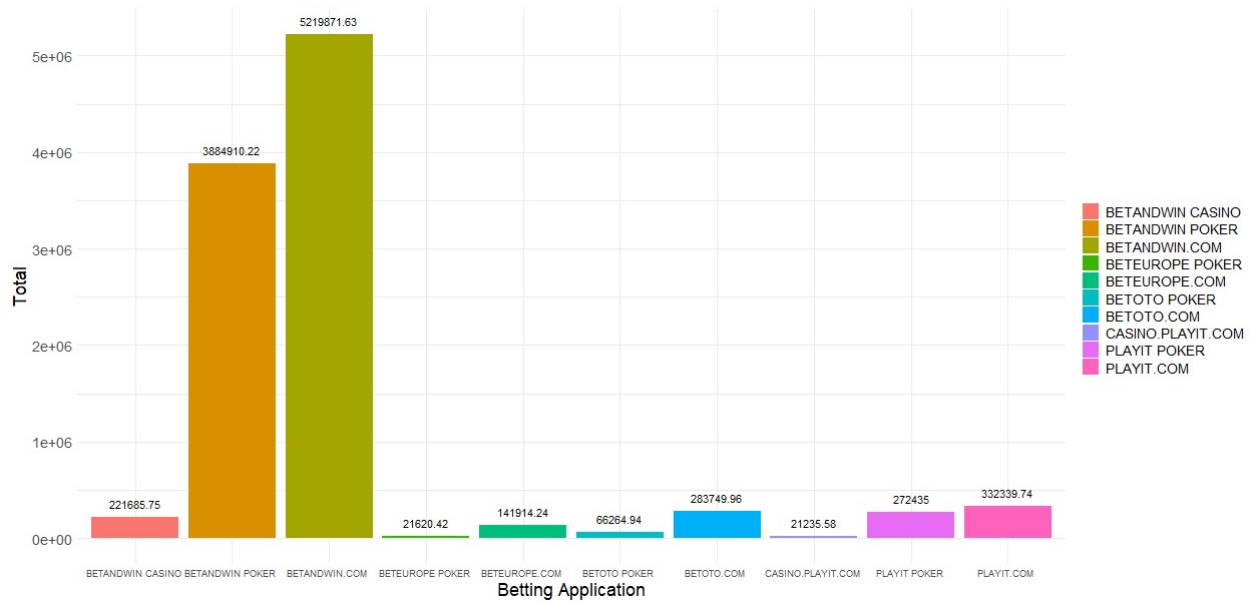


Figure 1: Total bets/users/amounts for each Country

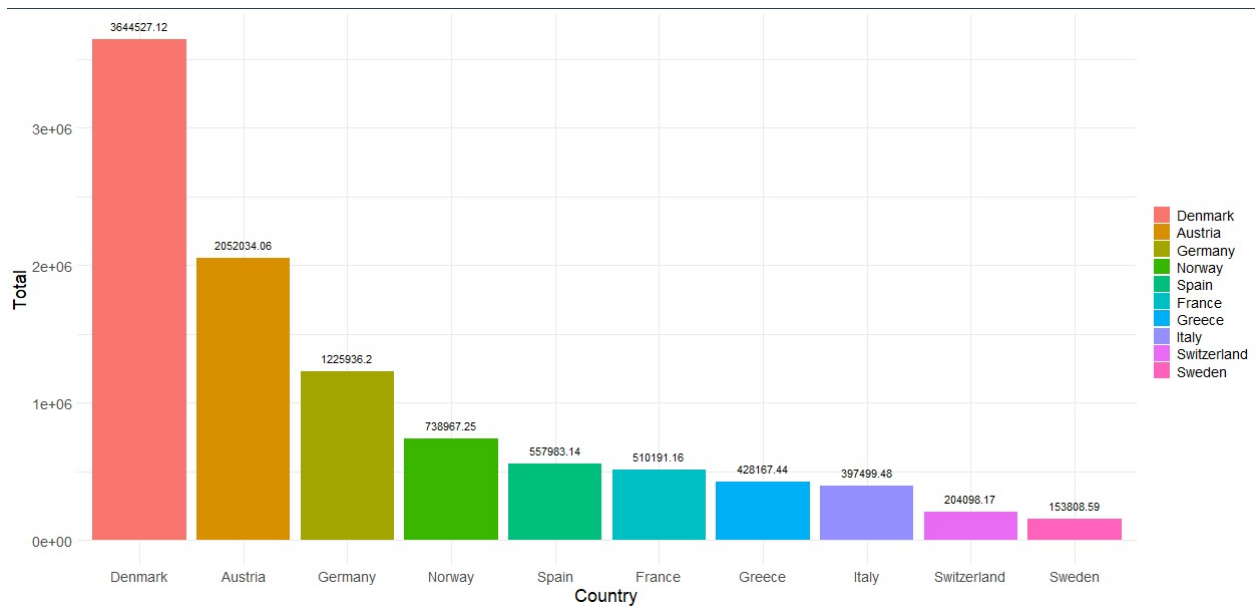


Figure 2: Total bets/users/amounts for each Betting Application



Figure 3: A world map of all countries to which the clients belong

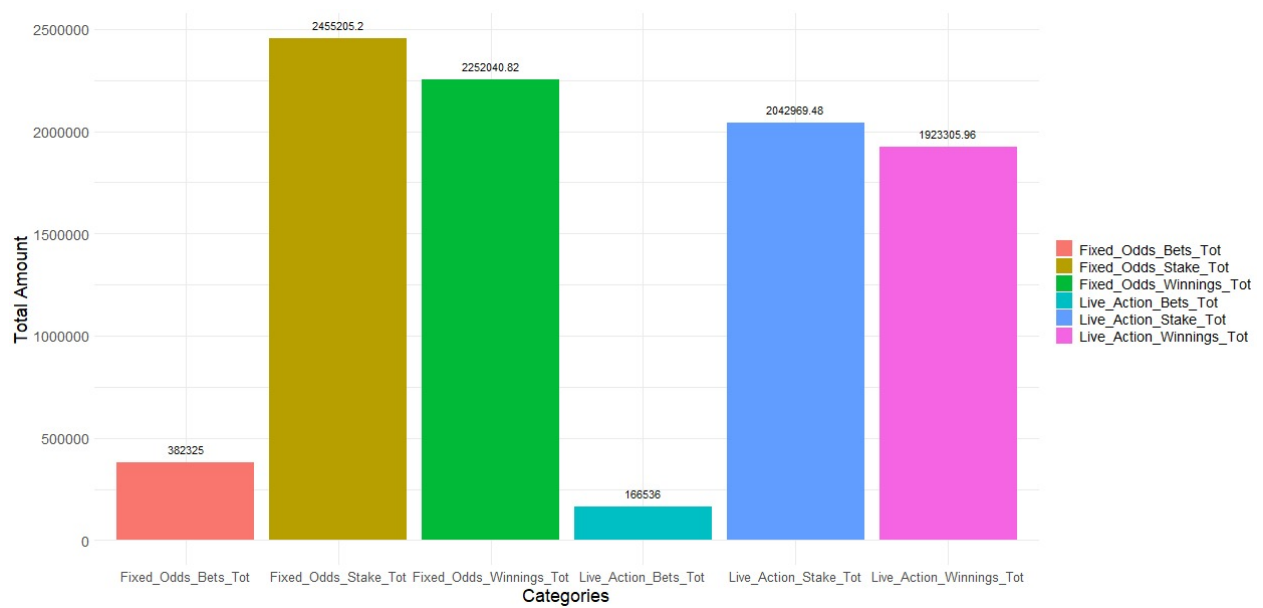


Figure 4: Total bets/stakes/winnings for Fixed Odds and Live Action product types.

period of the data provided.

#### **4. Shiny Application**

- A more detailed interactive analytics dashboard is available on the Shiny Application which has the following tabs:
- A comparison of the different betting applications
- A comparison of the countries
- A world map comparing the countries
- Comparing the Fixed Odds and Live Actions for each country

#### **5. Conclusion :**

- We have cleaned the data from BWin to create a detailed datamart describing the behaviour, demographics and stakes, winnings and bets for different products for each user with one row per user.
- We have created a dashboard to visualize the trends easily which and interactively allowing us to have a deeper insight of data and provide marketing analysis of the BWin users/.
- This Shiny dashboard is deployed on cloud and available publicly.