

Running head: UK Bank Customer Segmentation

UK Bank Customer Segmentation:

Jai Katariya

University of Colorado, Denver

<b>Sr.</b>	<b>Index</b>	<b>Page No.</b>
<b>1</b>	<b>Abstract</b>	<b>3</b>
<b>2</b>	<b>The Dataset</b>	<b>3</b>
<b>3</b>	<b>The Objective</b>	<b>3</b>
<b>4</b>	<b>EDA in R</b>	<b>4</b>
<b>5</b>	<b>Overview: UK Bank Customers</b>	<b>16</b>
<b>6</b>	<b>England</b>	<b>18</b>
<b>7</b>	<b>Scotland</b>	<b>20</b>
<b>8</b>	<b>Tableau Dual Filter – Scotland</b>	<b>22</b>
<b>9</b>	<b>Wales</b>	<b>23</b>
<b>10</b>	<b>Northern Ireland</b>	<b>25</b>
<b>11</b>	<b>Recommendations for Marketing Campiagns</b>	<b>27</b>
<b>12</b>	<b>References</b>	<b>28</b>

## Abstract

This PowerPoint explores the UK Bank Customers dataset with data visualization tool - Tableau. This PowerPoint starts with basic Map of UK to understand the dataset in Tableau with Basic nationality of the customers. It also tries to formulate business questions like “What is the target audience in different region” After this we try to answer the question with visuals and graphs with the help of Tableau interactive dashboard. We have 5 views to understand each variable is vital for our business question and perform customer segmentation by age, gender, job classification and region.

## The Dataset

It is a dummy dataset for an imaginary bank in UK. It's a new bank that is hoping to expand and run marketing campaigns at different regions depending upon the customer demographics like age, gender, job classification, balance, etc. It contains 4014 observation and 9 variables.

Sr. No	Variable	Data Type
1	Customer ID	Numeric
2	Name	String
3	Surname	String
4	Gender	String
5	Age	Numeric
6	Region	String
7	Job Classification	String
8	Date Joined	Date Time
9	Balance	Numeric

## The Objective

The objective here is to perform customer segmentation and capture the vital insights from interactive data visualization tool Tableau to help the marketing campaign in defining the target audience in each region.

### **Overview: Exploratory Data Analysis in R**

EDA or Exploratory Data Analysis is a process in which we generally obtain summary statistics, plots like histogram, etc. to graphically represent the relationships. EDA is a philosophy about how we dissect a data set, generally used at the beginning of the analysis to skim through our data.

#### **Top EDA functions in R**

- 1) Head() – This will return the first n observations of the dataset. Default is 6
- 2) Tail() – This will return the last n observations of the dataset. Default is 6
- 3) Summary() – This will return the summary of the statistics like min, median, max, quartiles for each column of the dataset. Count returns missing value if any.
- 4) Dim() – This will return the dimensions of the data frame like the number of rows and number of columns in the dataset. The output is a vector.
- 5) Hist() - This will return a bar plot in which each bar represents the frequency of cases for a range of values.
- 6) Str() – returns the classes of each of the columns whether it is numerical or categorical data.
- 7) Names() – This will return the column headers.

#### **Missing Value**

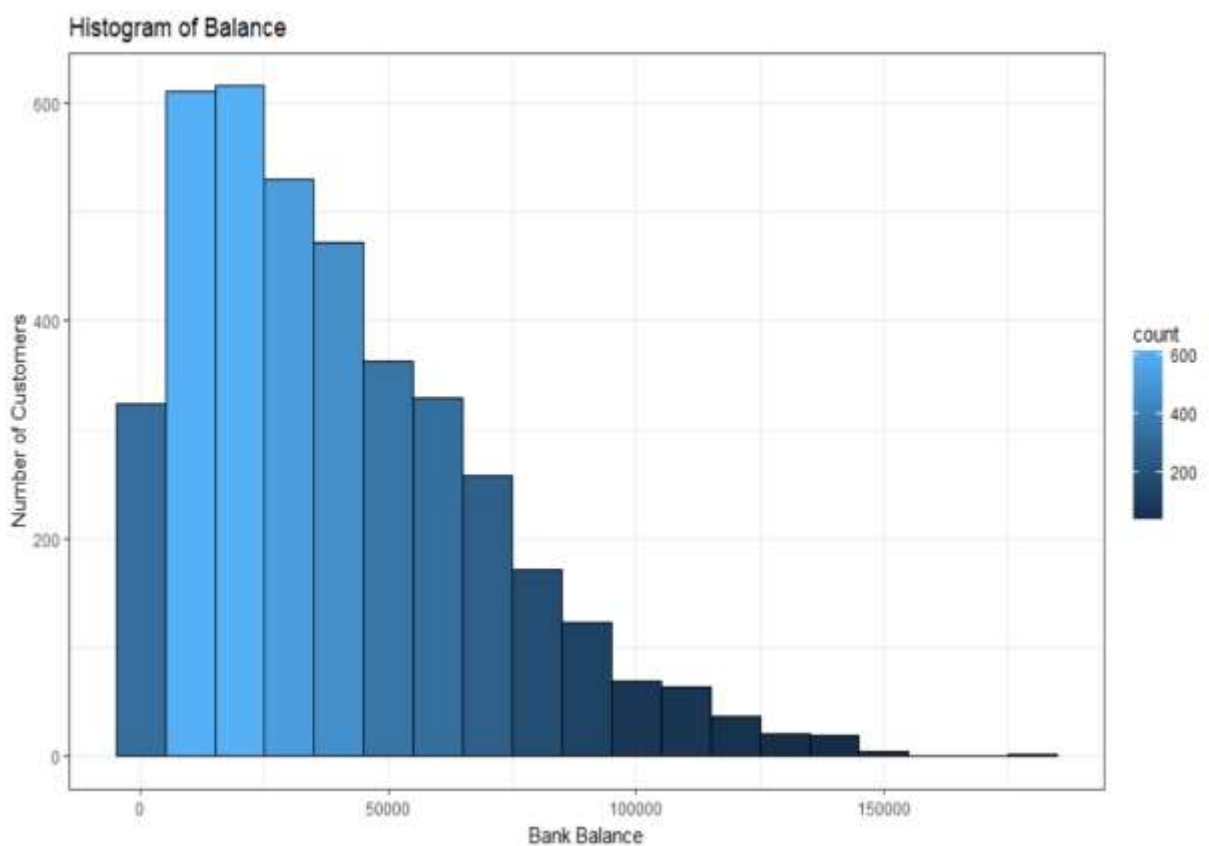
Before EDA, I checked for missing values. To do this I applied Amelia library. The dataset was clean and I found no missing values.

## Univariate Analysis

Here, I took variables and plotted the frequency distribution graph.

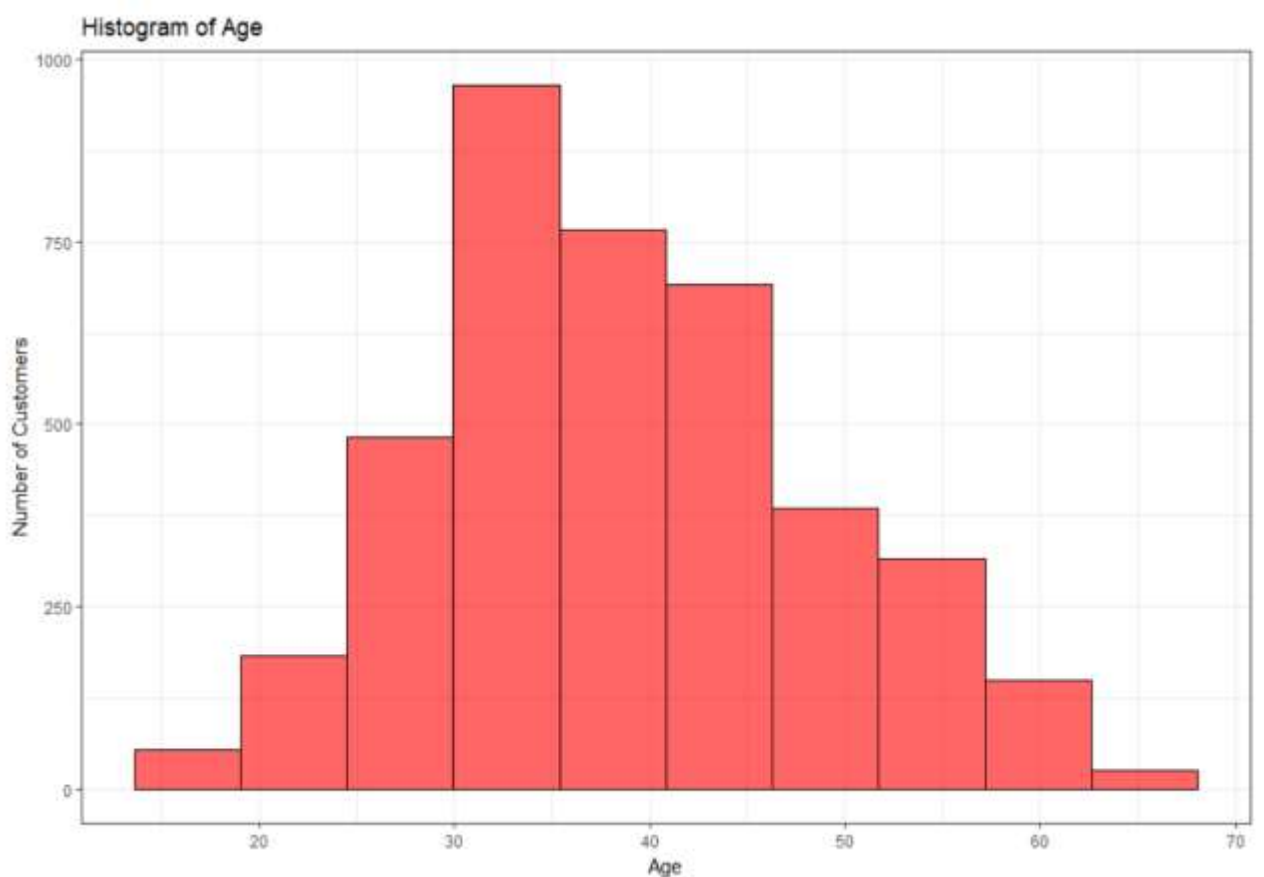
### 1. Histogram of Balance

The variable “Balance” refers to Bank Balance of the customer. As noticed, we can see majority of the customers have a bank balance of within the range of \$10000 to \$30000. We can also see that there is an outlier at around \$180K as well.



## 2. Histogram of Age

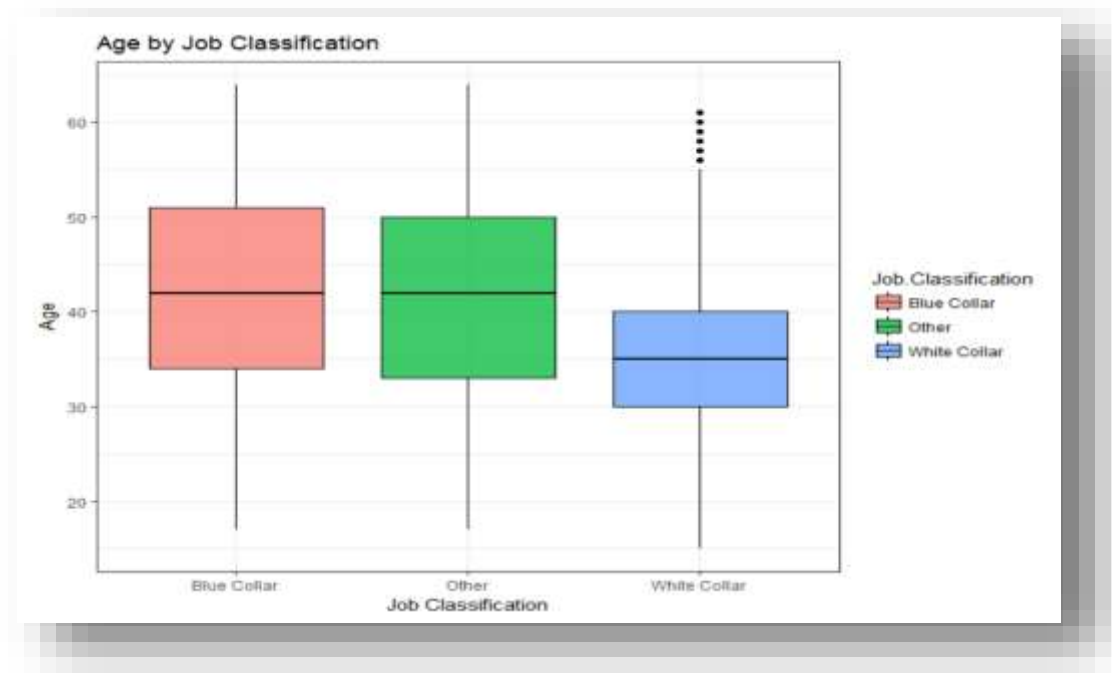
In the below graph the “Age” variable represents the age of the customer. Here, we can notice that it is fairly normally distributed with a peak around 30 to 35 years old customers. Majority of the customers are in the range of 25 to 50. This is an important feature for the marketing department as this age group is a potential target audience.



## Bivariate Analysis

### 1. Boxplot of Age by Job Classification

The variable “Job Classification” represents the customer’s job. And it is a categorical variable and thus we make a box-plot.

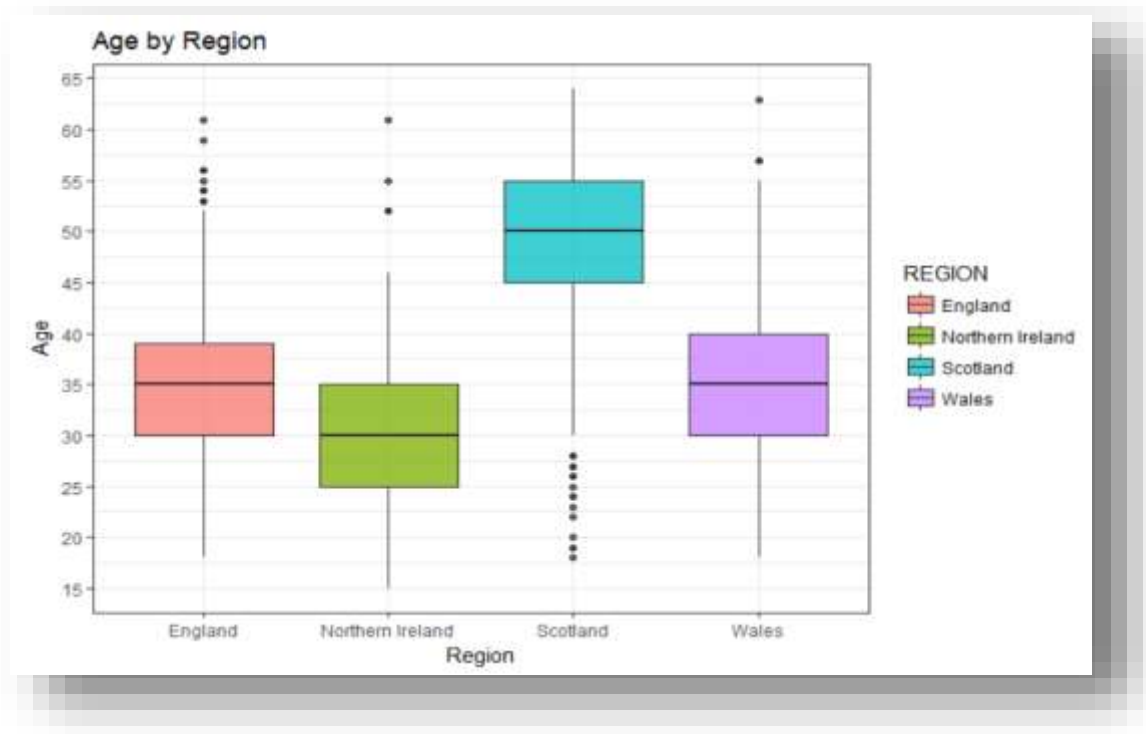


### Insights

- Here, we notice that Blue collar and other Job Classification are in the Age group of 33 to 52 with a median of 42 years old
- Whereas, the “White Collar” group has a younger audience ranging between only 30 to 40 years and median of 35.
- Also, we can see that there are outliers in White Collar between 50 to 60 years old group. This is an interesting insight.

## 2. Boxplot of Age by Region

There are four regions in the dataset - England, Northern Ireland, Scotland and Wales. This plot is very important as the marketing department will know the age by geography.



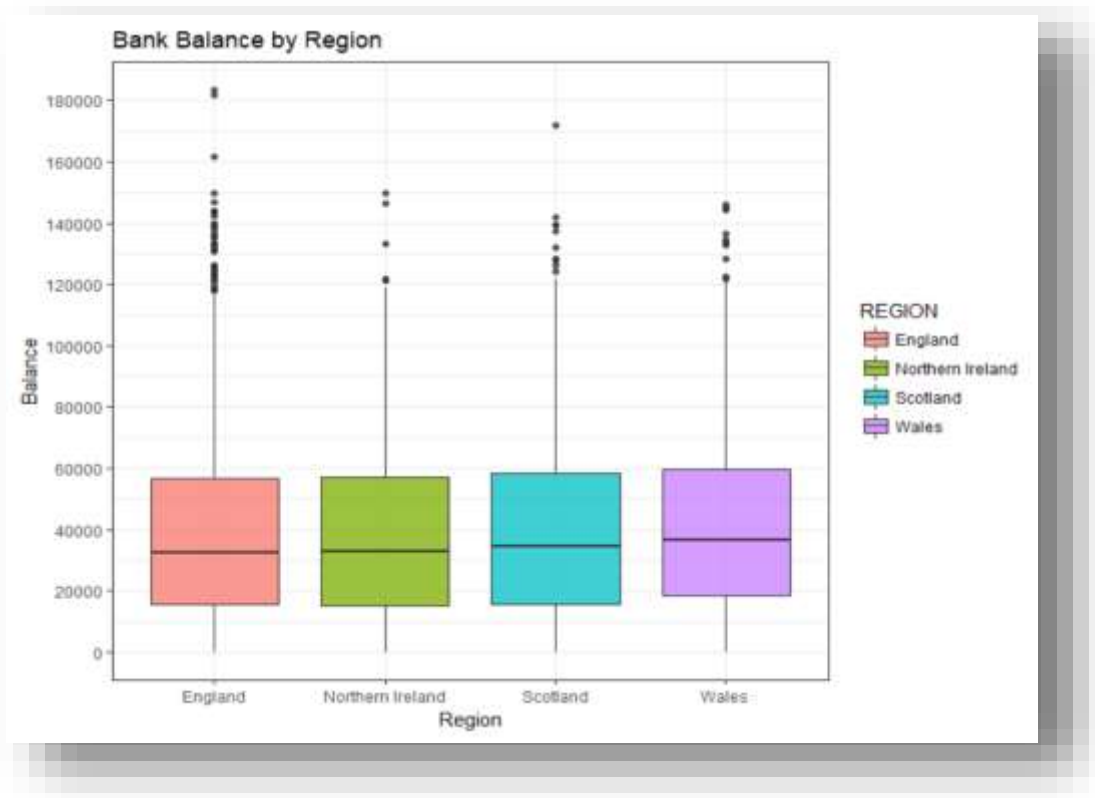
### Insights

- Scotland stands out the most. It represents the older customers in overall the region with the range of 45 to 55 and median of 50. But it also contains a long lower end tail between the 20's and 30's
- Northern Ireland – On the other hand, Northern Ireland has a younger generation group that lies between 25 and 35.
- Wales looks normal with the age range of 30 and 40
- England – as we noticed over 2000 customers are from England. This shifts the age distribution towards England. England has an age range between 30 and 40. Also has a huge tail at the older age above 50 years.



### 3. Bank Balance by Region

This graph is important as it shows overall the customers are having similar distribution that ranges between \$15k and \$60k.

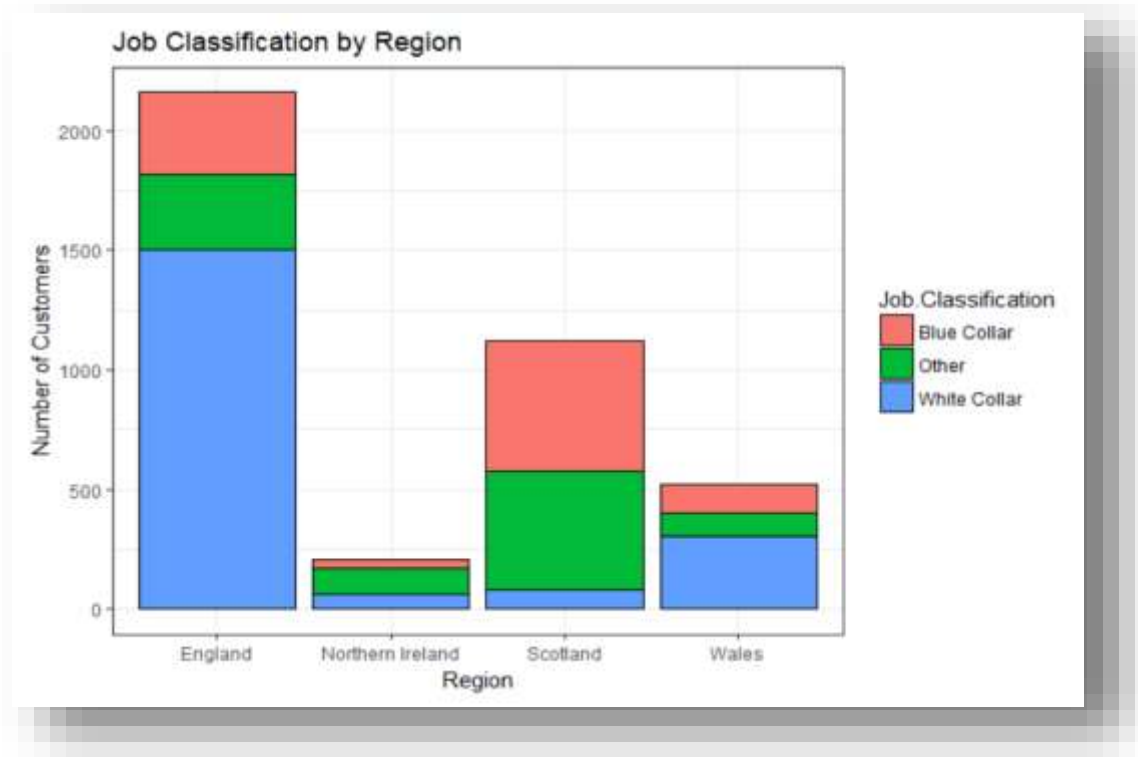


#### Insights

- England has a long higher tail showing the higher balance customers are in that country.
- Wales has a higher boxplot stating that overall customers in that region are likely to have a higher bank balance. This could help the company while selling their products like Debit cards that largely depend on the bank balance.

#### 4. Job classification by Region.

This graph represents the customers job type pattern in each region.



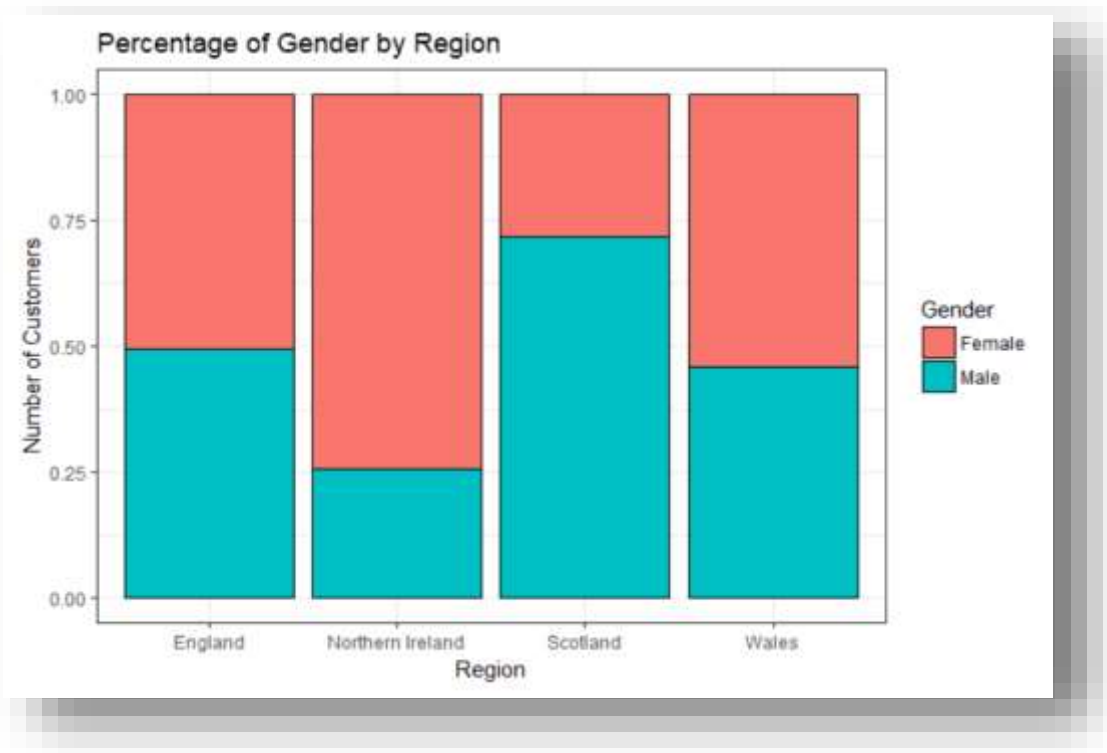
#### Insights

- Around 70% of the customers are white collar in England. This might be something related to London that is the finance hub and majority of people there are white collared.
- On the other hand, Scotland has no or very less white collar customers.
- Around 80% people are white collared in Wales

This is a very important finding. The marketing group can target the England markets keeping the white collar audience as their major segment. On the other hand, Scotland needs a different marketing approach.

## 5. Gender by Region

The below graph represents the gender demography in each region. Gender is a vital variable in marketing and targeting the audience.



### Insights

- As we can see, England and Wales have similar customer demographics. They have around 50-50 for males and females.
- Northern Ireland is interesting as 75% of the customers are females.
- Scotland on the other hand is only 25% for females and is a male dominating region.

## 6. Age by Gender

Here we see the overall banks customer's age and gender.



### Insights

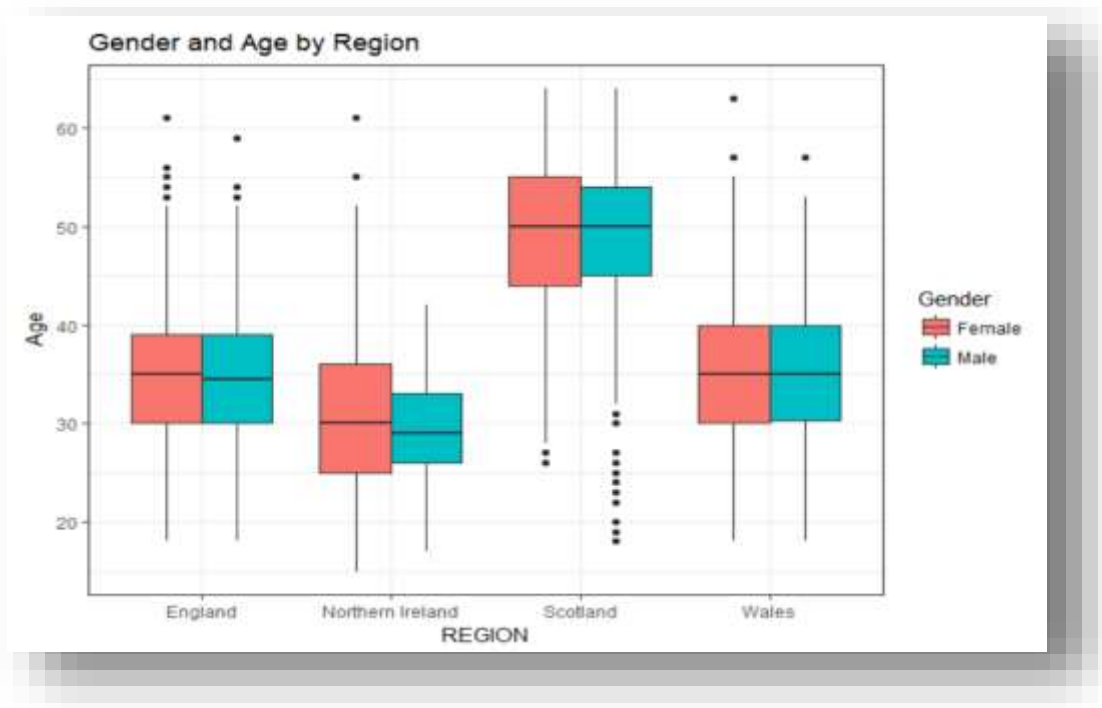
- Female segment is ranging between 30s and 40s
- Female have outliers around 60s as well.
- Male segment has a bigger range from early 30s to late 40s.

## Multivariate Analysis

### 1. Gender and Age by Region

As we have noticed the region is an important variable we are incorporating it in our analysis.

Below is the graph represents the customers age and gender in a particular region.

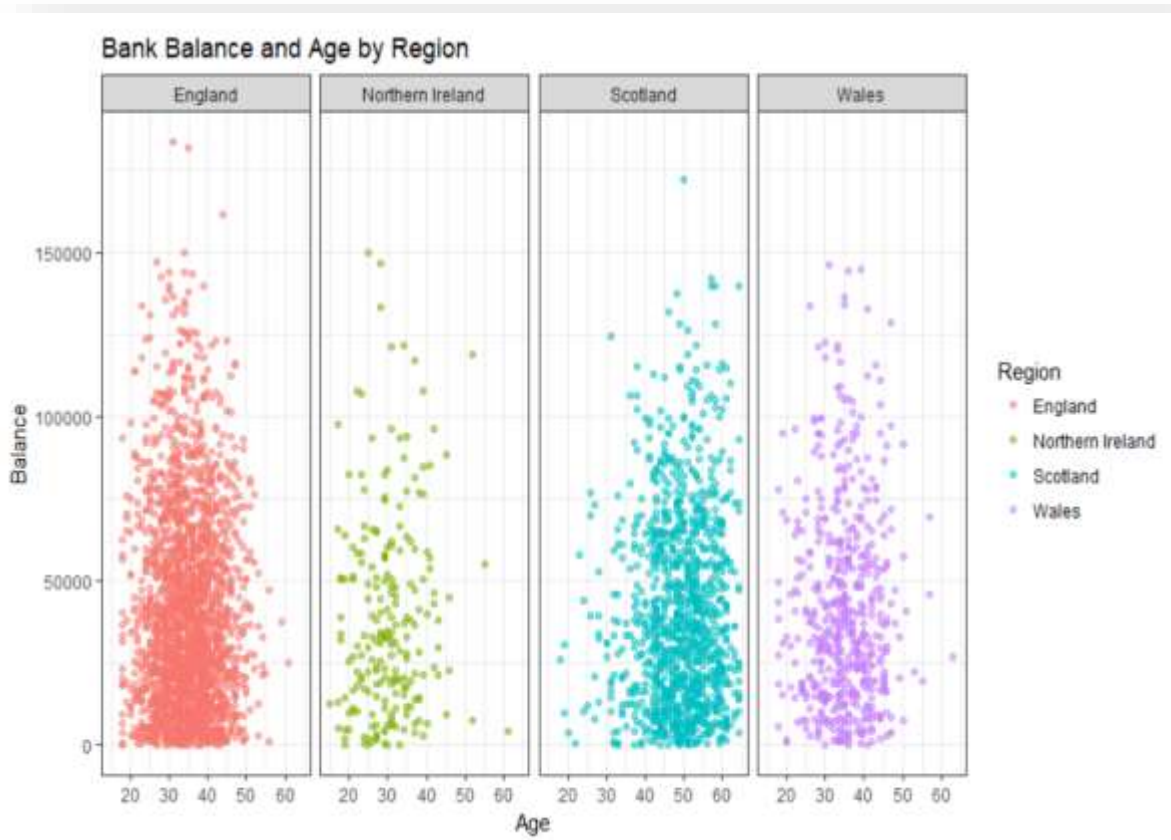


### Insights

- England and Wales have similar boxplots for the age component as well as the age component.
- Northern Ireland has more females with a bigger box representing that range from 25 to late 30s whereas males range from 27 to early 30s.
- Scotland has an outstanding box plot among the four regions. The age ranges from late 40s to late 50s but it has a long tail for the younger males around the age of 20s and 30s.

## 2. Bank Balance and Age by Region

The below graph represents age distribution for bank balance in particular region.

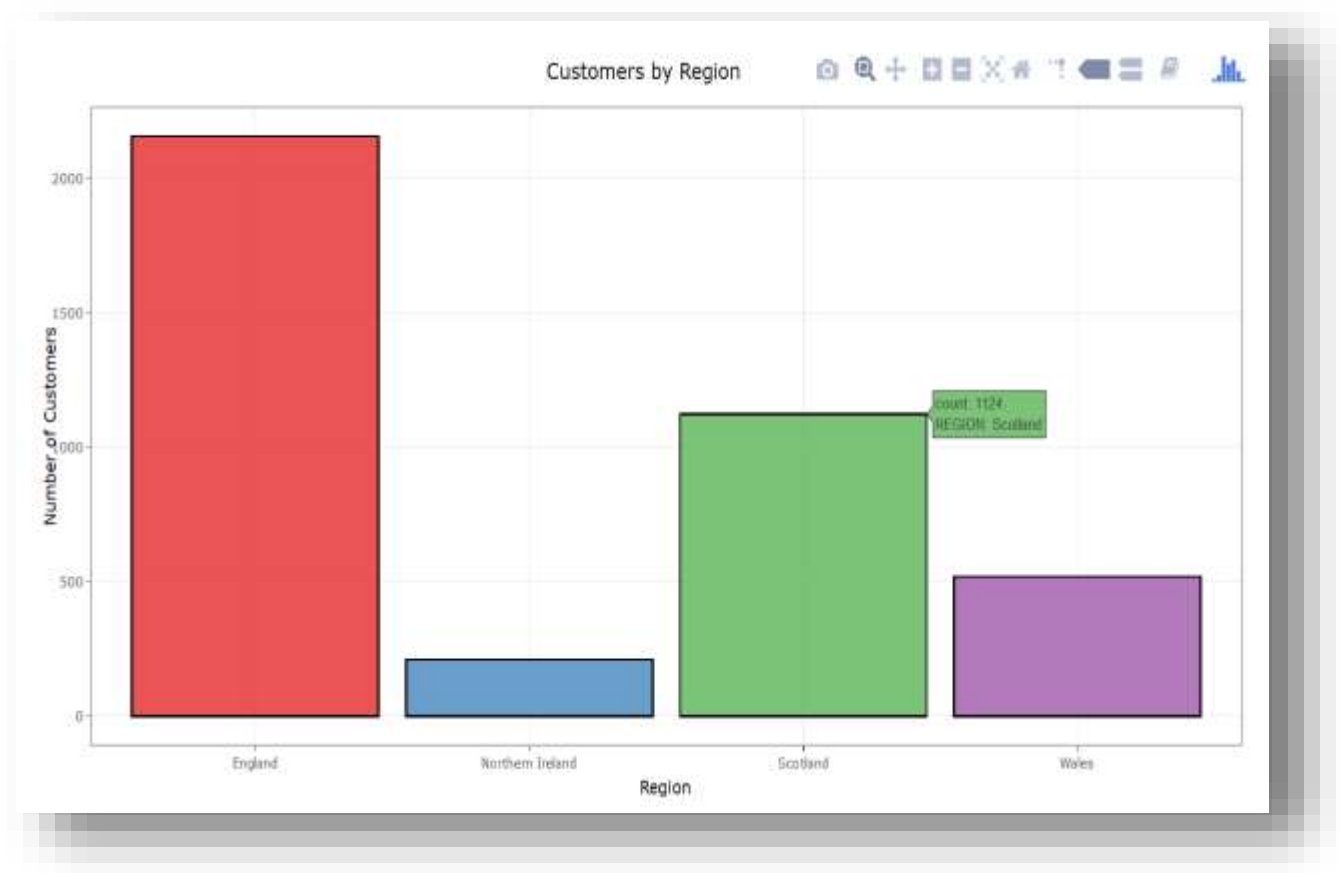


### Insights

- England and Wales are condensed around the age of 20 to 40 and have a bank balance between \$5K and \$65K.
- Northern Ireland has younger customers more below late 20s and the bank balance has little concentration below \$50K.
- Scotland has the older generations ranging from early 40s to late 50s. The bank balance ranges between \$5K up to \$75K.

## Data Visualization with PLOTLY

Lastly, I wanted to check whether we can have **some interactive plots in R**. Therefore, I took the customers data and segmented it by Regions, which are countries in the dataset. I created a bar plot in GGLOT and recreated that with the help of PLOTLY Library to get an interactive plot. The below graph represents the customer segmentation by region and the number of customers in that region when we hover over the plot.

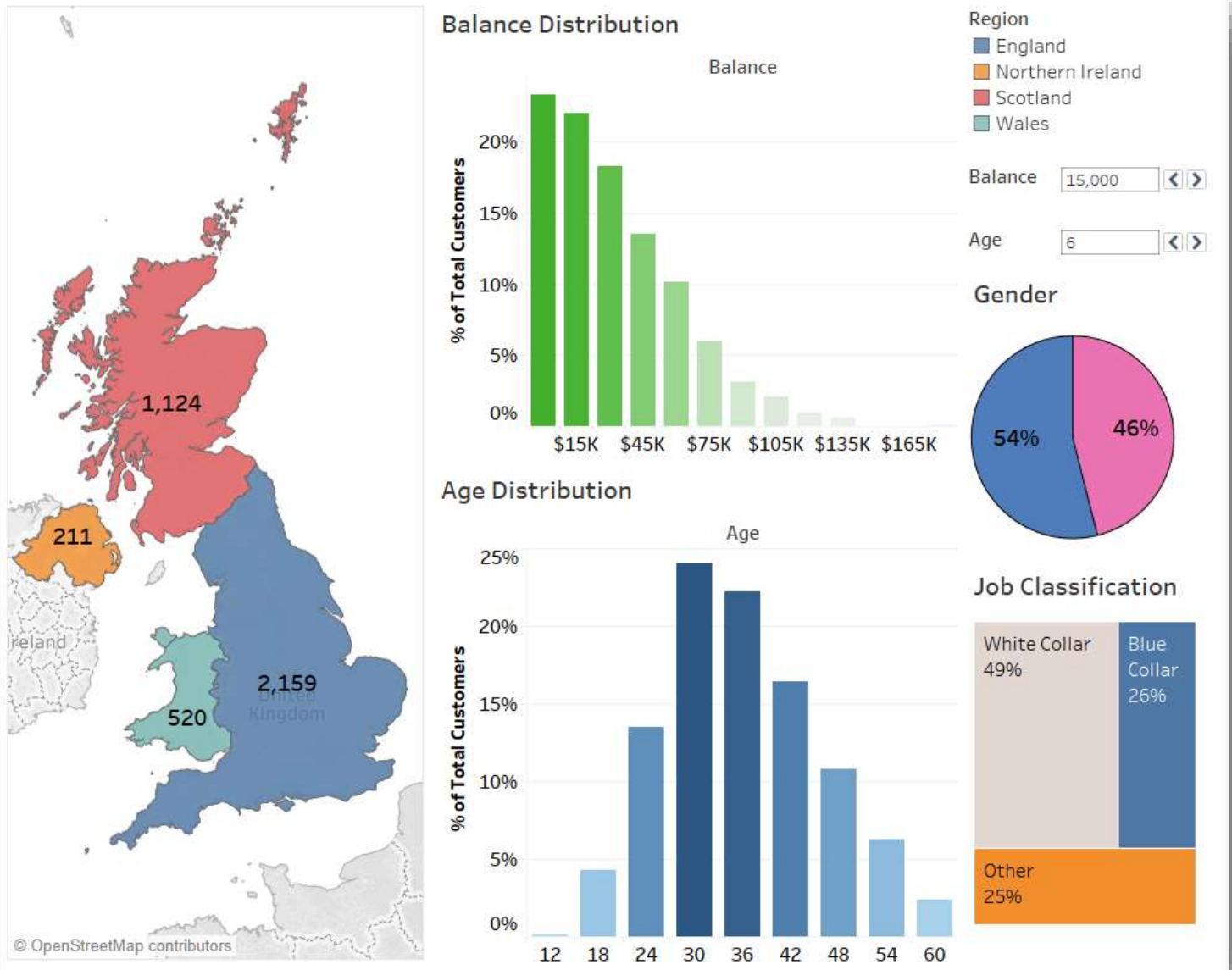


### Insights

- 50% of the data is from England.
- 25% from Scotland.
- Northern Ireland has the lowest number of customers.

### Overview: UK Bank Customers

In slide 3 the dashboard has 5 views, helping us to get vital information and insights about the existing customers. It also helps the marketing department to lay the foundations for the target audience and next marketing campaign.



### Overview



- From the map of UK we can notice that majority of the bank's customers are from England.
- Majority of customers' bank balance is below \$50K mark.
- Average range for the age of the customers is 30 to 44.
- Gender is almost equal.
- Around 50% of customers are white collared, showing England been major contributor and London being the hub for white collar can be a reasonable explanation.

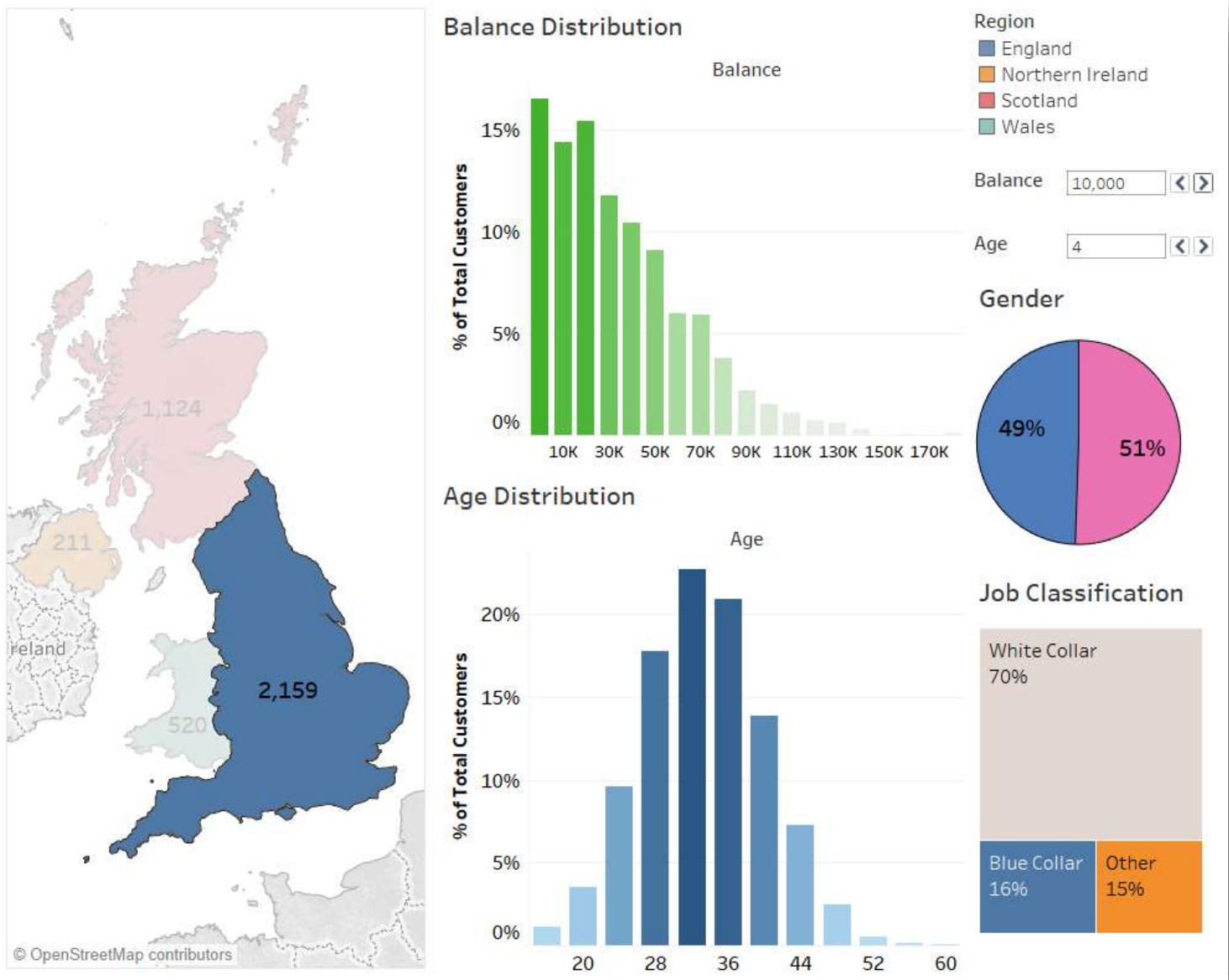
The above information sets a general overview of the current customer type. But marketing campaigns aren't similar for all the regions. Therefore, with the help of this PowerPoint and Tableau's interactive strengths, we can dissect the customer segment and help the marketing and sales department to target a particular type of customer in each region.

### Region Analysis

Here, I'll take each region and see if it is different from our general customer type.

#### 3. England

In slide 4, I have selected England and due to the interactive feature of Tableau other views highlights England's contribution in them.

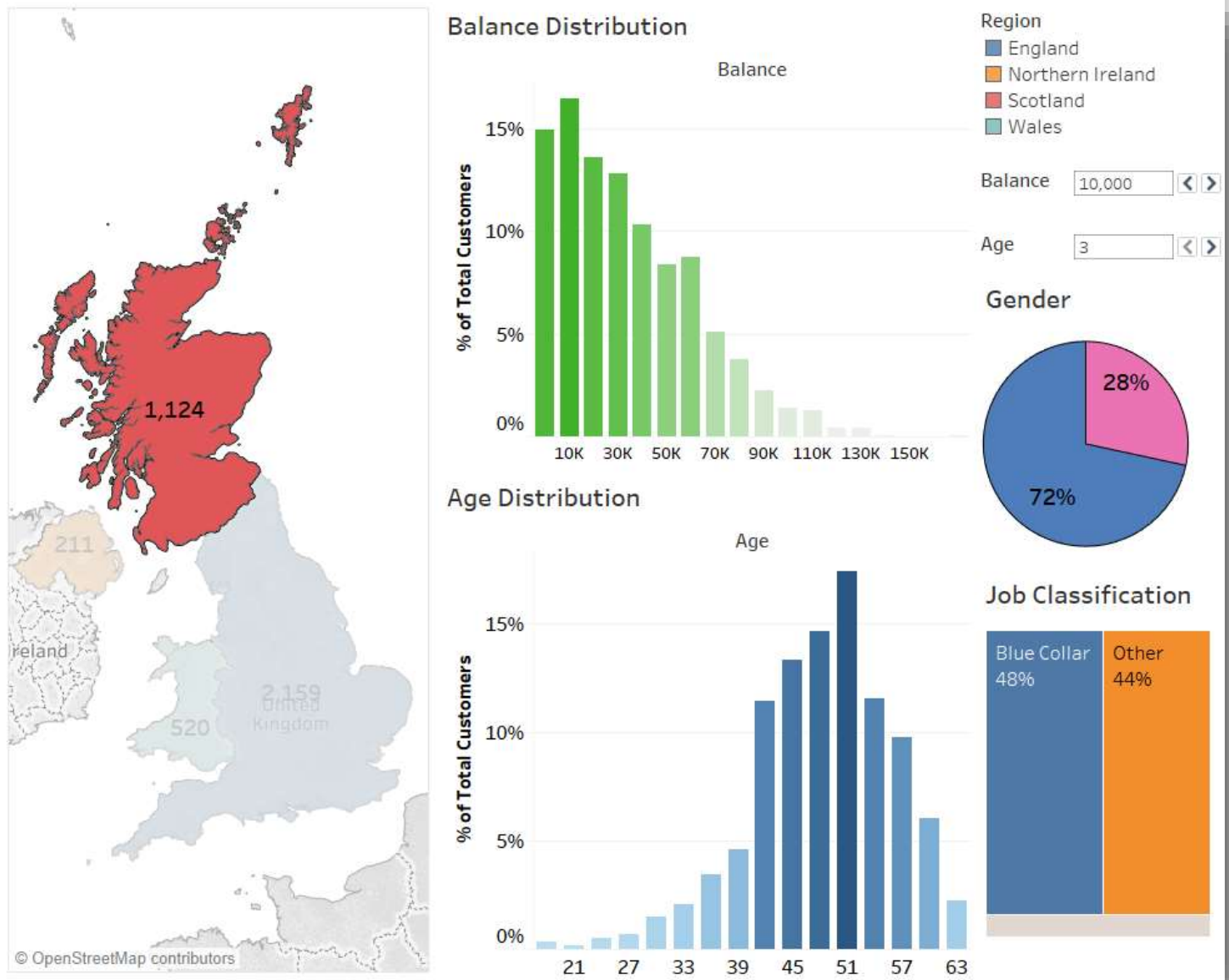


**Insights**

- The balance histogram is on an average similar to the overall UK balance.
- Age is an interesting feature as there is a huge drop after the age group of 35-37, which says England has younger customers compared to UK customers. This is something the marketing team should keep in mind while planning their strategies in England.
- 70% of the England's customers are White Collar. This is important because it means the marketing team needs to target white collar people as they are the potential customers.
- Gender is equal.

#### 4. Scotland

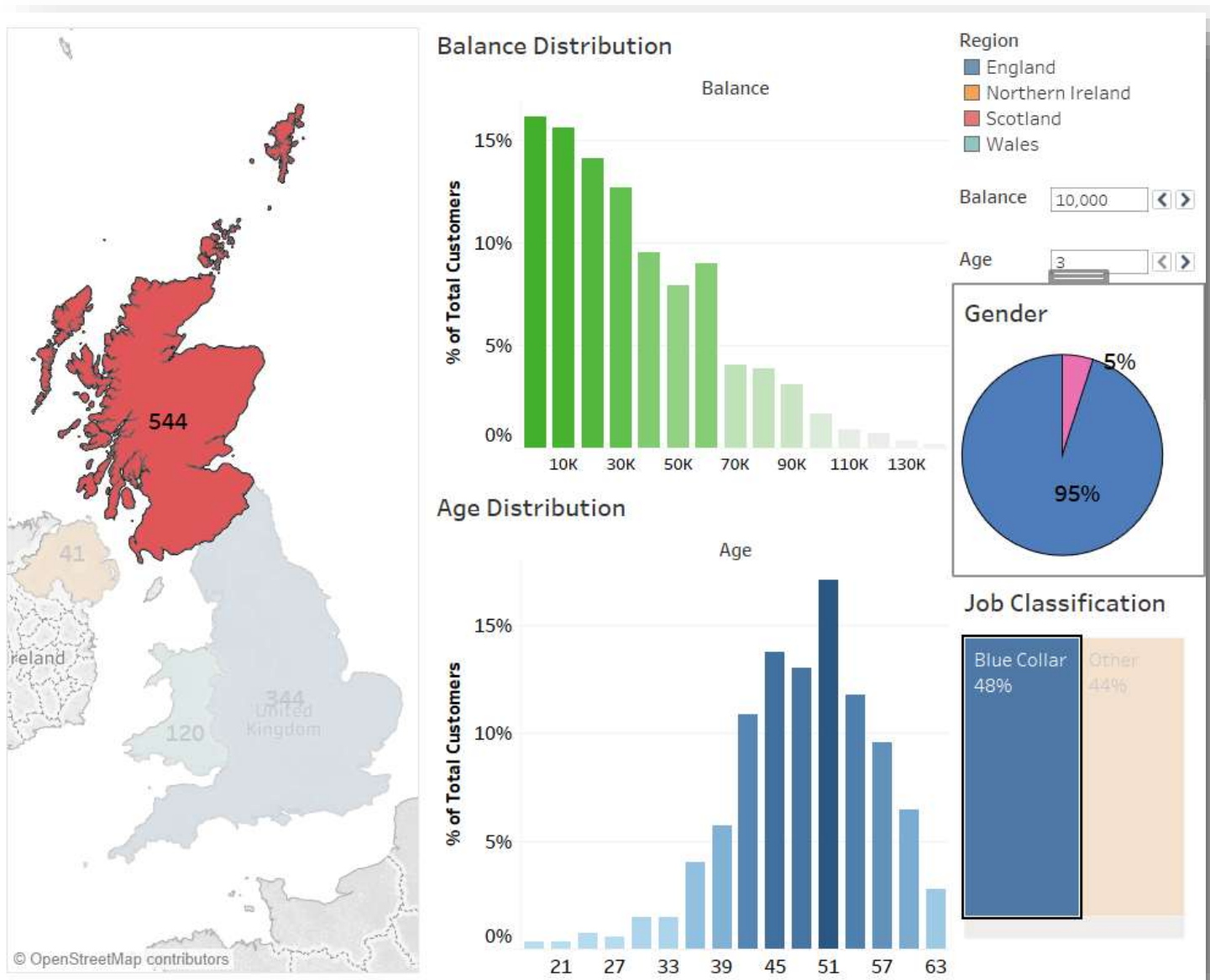
In slide 5, I have selected Scotland and due to the interactive feature of Tableau, other views highlights Scotland's contribution in them.



**Insights**

- The Bank Balance is slightly different than the UK's overall bank balance. It is more like steps and the range has increased starting from \$5K to \$70K.
- Age is the most vital feature in Scotland, catering to older customers with an age ranging from 45 to 55. The marketing and sales team needs to keep the age feature in mind when targeting Scotland and prepare ad campaigns accordingly.
- Scotland is a male-dominating country with respect to the existing customers since 72% of customers in Scotland are males.
- Lastly, Blue collar and others are dominating this region with 7% of white collar.

## Tableau Dual Filter

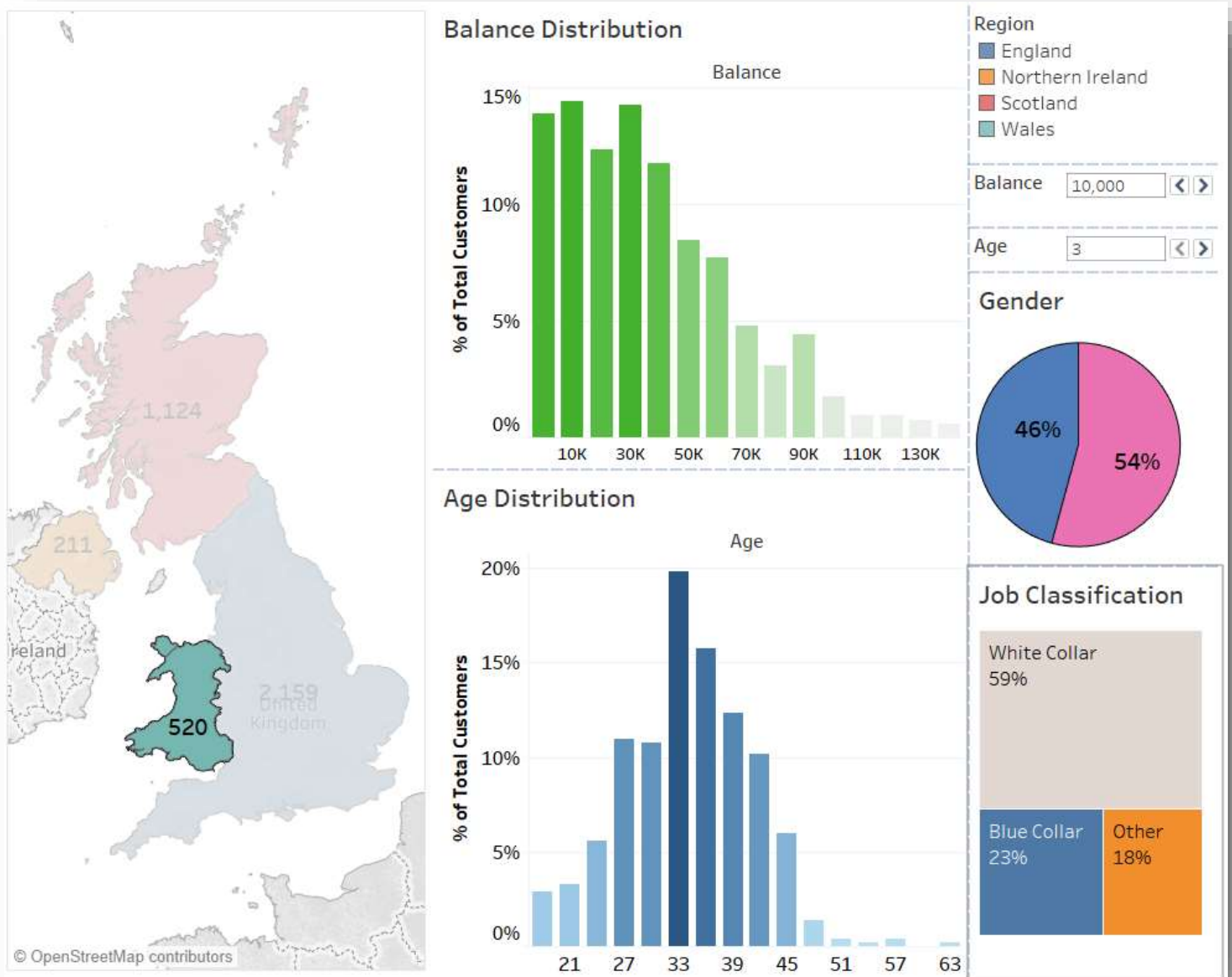


## Insights

- When we dissect the customer by Scotland and Blue collar, we get **95% of male dominated customer.**

## 5. Wales

In slide 6, I have selected Wales and due to the interactive feature of Tableau, other views highlights Wales's contribution in them.



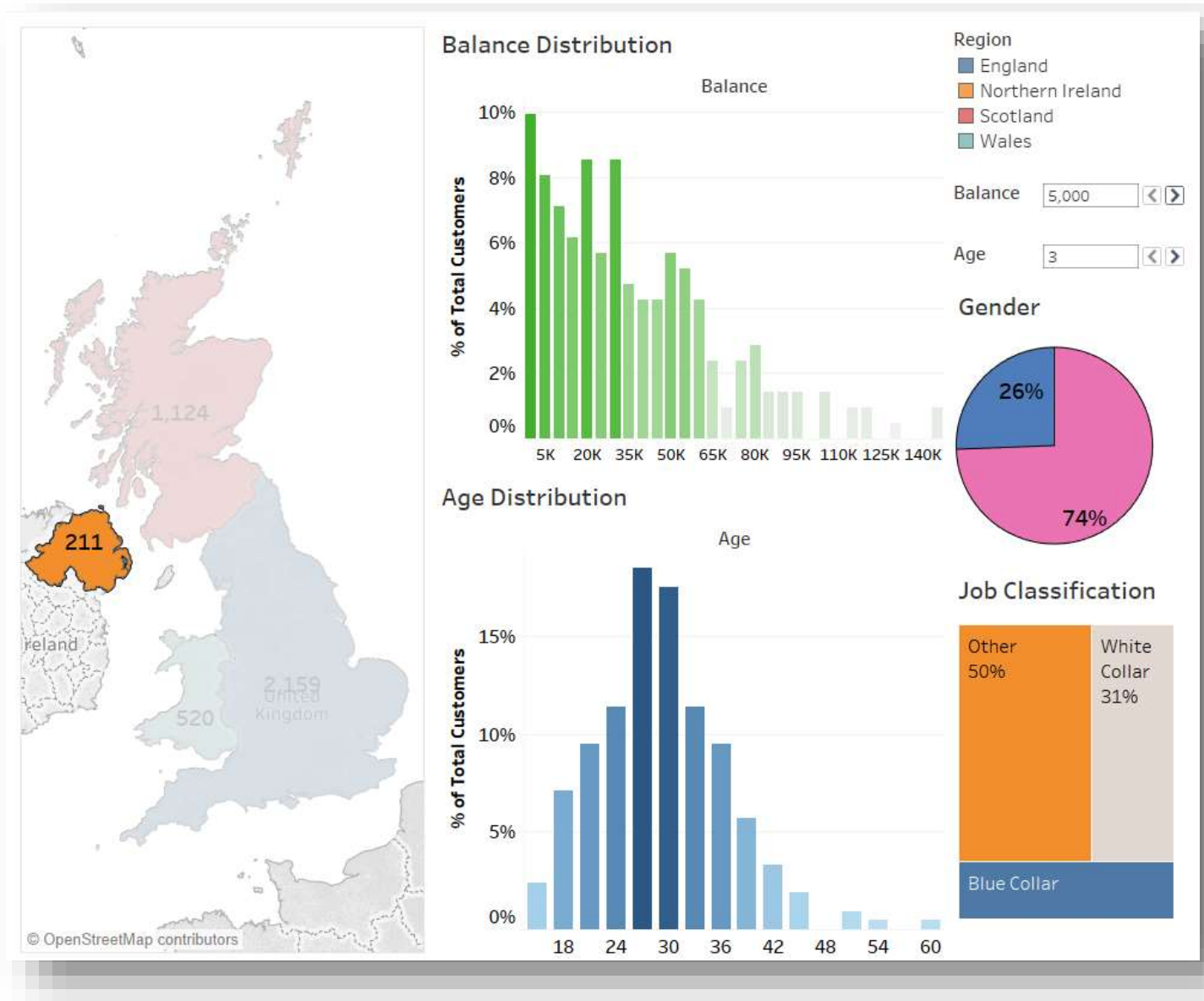
**Insights**

- The Bank Balance has been uniformed till \$45K and then slides downwards.
- In Wales the age distribution is also quite uniformed ranging from 25 to 45, with a huge peak at 34, almost acting as an outlier.
- Gender is almost equal.
- Around 60% of customers in Wales, are white collared.



## 6. Northern Ireland

In slide 7, I have selected Northern Ireland and due to the interactive feature of Tableau other views highlights Northern Ireland's contribution in them.



**Insights**

- The Bank Balance is different compared to UK's overall bank balance. It has high bumps at \$20K and \$30K. This allows the promotional team to have some insights like giving benefits for keeping a balance around \$20K to \$30K, which attracts the younger generation to be the potential customer.
- Age is the most vital feature in Northern Ireland as represent the younger customers ranging from 24 to 33. Although it is a small market but if the marketing department targets the younger generation, it can be one of the potential markets in the upcoming years. Also there is a huge gap between 47 and 52 age group.
- Job Classification is another good variable suggesting that unlike other regions it has lower white and blue-collar customers. Here, Job Classification represents 50% as others.
- Gender is 74% females, which can totally help the marketing department.

## Recommendations for Marketing Campaigns in each region

In slide 8, I have given few recommendations for the marketing campaigns to target the audience.

Below are the few recommendations according to each region.

### → England

- England has 70% of the customers as white collar. Thus, the marketing department and sales department should target white collar and give offers **like investment plans and even long term mutual fund investments** that this specific audience who are in their early 30s can relate to.

### → Scotland

- Scotland is dominated by old males with the job classification as blue collar and others. Thus, the marketing team should focus on old men in Scotland with good deals in **Retirement plans especially for blue collar type of customers.**

### → Wales

- Wales has more of mid-30s, white collared customers. Thus, the marketing team should target them with benefits related to white collar jobs like **Investment plans and even Long Term Mutual Funds.**

### → Northern Ireland

- Northern Ireland has more young women. Thus, the marketing teams should target young women with benefits of having balances around \$30K and can offer **Educational loans for women.**

## Reference

<https://www.superdatascience.com/>