Katarina Milivojević

- Data Analyst with four years of working experience with clients from various countries, and fields such as human resources, marketing, social psychology and medicine. Particularly skilled with Excel, SPSS, Python, R, and Tableau.

"If you torture the data long enough, it will confess." — Ronald H. Coase

Projects

- **Covid-19 testing rate**
  - visualization with ggplot2 in R
- **Likelihood of woman having a third birth**
  - survival analysis in R
- **Post-millennials' willingness to purchase products with GI**
  - structural equation modeling in SPSS AMOS
- **Effects of the type of Instagram endorser on the purchasing intentions of their followers**
  - survey experiment data analysis
- **Profile of the UK bank customers in different regions**
  - Tableau story

Covid-19 testing rate

How does testing rate changes in European countries over weeks, while taking positivity rate into account?

Data set: downloaded from European Centre for Disease Prevention and Control: https://www.ecdc.europa.eu/en/publications-data/covid-19-testing
It was about testing for COVID-19 by week and by country, with 1099 observations.

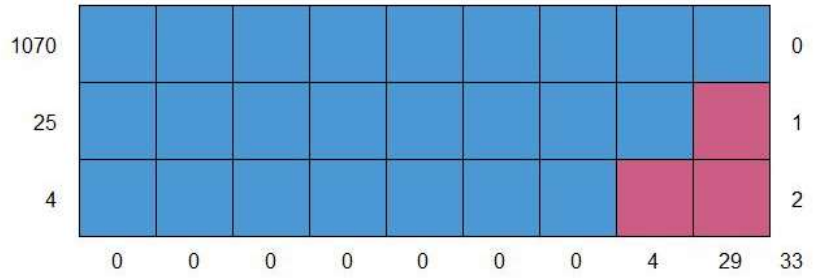
	A	B	C	D	E	F	G	H	I
1	country	country_code	year_week	new_cases	tests_done	population	testing_rate	positivity_rate	testing_data_source
2	Austria	AT	2020-W15	2041	12339	8858775	139.2856236	16.54104871	Manual webscraping
3	Austria	AT	2020-W16	855	58488	8858775	660.2267243	1.461838326	Manual webscraping
4	Austria	AT	2020-W17	472	33443	8858775	377.5126922	1.411356637	Manual webscraping
5	Austria	AT	2020-W18	336	26598	8858775	300.2446727	1.263252876	Country website
6	Austria	AT	2020-W19	307	42153	8858775	475.8332839	0.728299291	Country website
7	Austria	AT	2020-W20	363	46001	8858775	519.2704409	0.78911328	Country website
8	Austria	AT	2020-W21	267	39348	8858775	444.1697639	0.678560537	Country website
9	Austria	AT	2020-W22	231	46677	8858775	526.9012928	0.494890417	Country website
10	Austria	AT	2020-W23	184	41063	8858775	463.5290997	0.448091956	Country website

Variable “positivity_rate” was obtained by dividing “new_cases” by “tests_done” * 100.

Covid-19 testing rate

How does testing rate changes in European countries over weeks, while taking positivity rate into account?

Data cleaning:
29 missing values in “positivity_rate” and 4 in “testing_data_source”
Monotone missing data pattern:



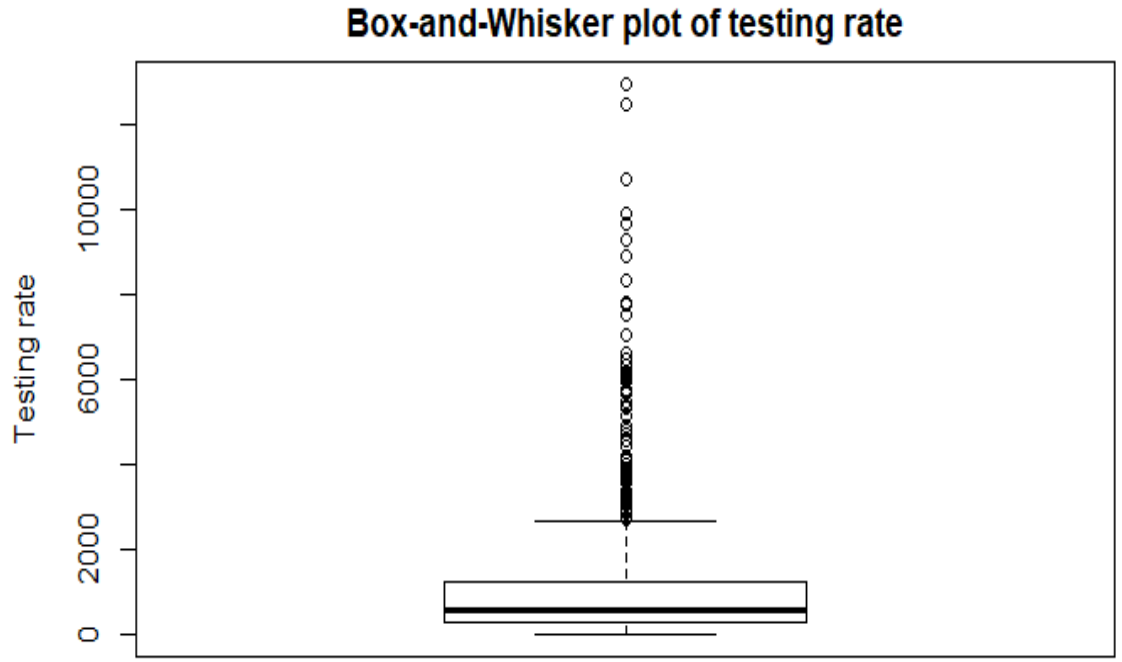
Rows with missing values were dropped, as they were all considering only first several weeks of the year, and the cells needed for calculating positivity rate were almost always zero.
Variable “year_week” was transformed by keeping only the week number and transforming it to numerical value.

Covid-19 testing rate

How does testing rate changes in European countries over weeks, while taking positivity rate into account?

Descriptive characteristics of variable “testing_rate”

- mean: 1011.25
- median: 579.91
- variance: 1 939 297
- minimum: .05
- maximum: 12 947.02
- 1st quartile: 271.13
- 3rd quartile: 1 233.90



Covid-19 testing rate

How does testing rate changes in European countries over weeks, while taking positivity rate into account?

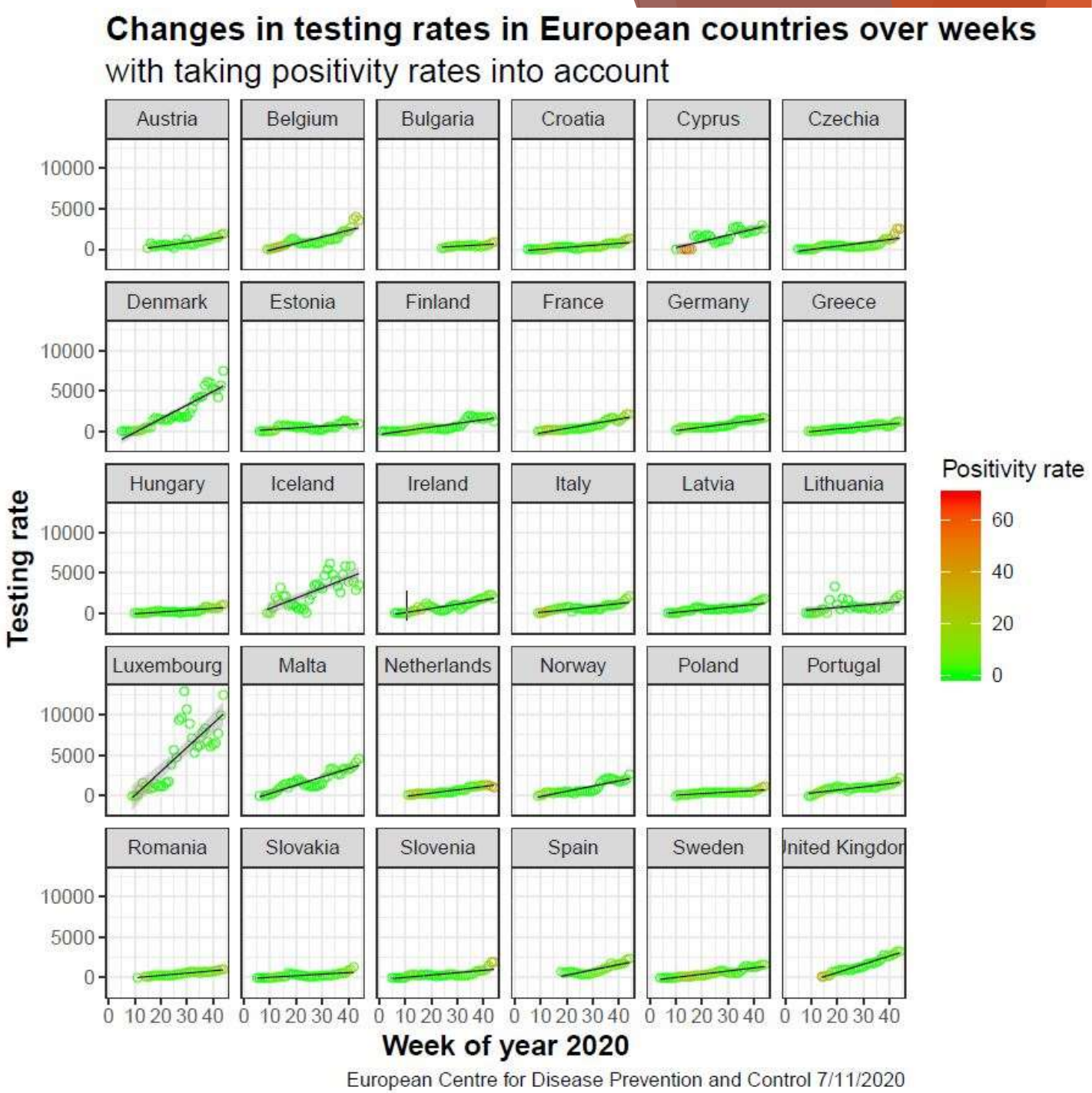
```
plot = ggplot(df, aes(x=df$year_week, y=df$testing_rate))
+ geom_point(shape = 1, aes(color = df$positivity_rate))
+ geom_smooth(method = lm, size = 0.3, color = 'black')
+ scale_color_gradient(low='green', high='red')
+ theme_bw() + facet_wrap(~df$country)

plot = plot + labs (
  title = 'Changes in testing rates in European countries over weeks',
  subtitle = 'with taking positivity rates into account',
  caption = 'European Centre for Disease Prevention and Control 7/11/2020',
  color = 'Positivity rate',
  x = 'Week of year 2020',
  y = 'Testing rate')

plot = plot + theme(plot.title = element_text(size = 15, face = 'bold'),
                    plot.subtitle = element_text(size = 15),
                    axis.title.x = element_text(size = 13, face = 'bold'),
                    axis.title.y = element_text(size = 13, face = 'bold'))

plot

pdf('plot.pdf')
print(plot)
dev.off()
```



Survival analysis

What is the effect of gender of first two children on the likelihood of a woman having the third birth?

Data set: Time to third birth data of the Medical Birth Registry of Norway, containing records of all births in Norway since 1967.
Downloaded from the selection of 16 116 cases from the initial data set (http://folk.uio.no/borgan/abg-2008/data/third_births.txt).

Variables:

- ***age*** (age of mother at first birth in years),
- ***spacing*** (time between first and second birth in days),
- ***sibs*** (genders of the first two children, 1 = boy, boy, 2 = girl, girl, 3 = boy, girl, 4 = girl, boy),
- ***time*** (time from second birth to third birth or censoring in days),
- ***status*** (censoring indicator, 0 = censored, 1 = birth).

Right-censoring data.

Survival analysis

What is the effect of gender of first two children on the likelihood of a woman having the third birth?

Data preprocessing

- No missing data
- Adding variables:
  - ***ID*** (unique identifier)
  - ***spacing_months*** (time between 1st and 2nd birth in months; *spacing* divided with 30.5)
  - ***time_months*** (time from 2nd to 3rd birth or censoring in months; *time* divided with 30.5)

Descriptive characteristics of the sample

age, M = 23.32, SD = 2.20
spacing_months, M = 32.61, SD = 14.02
10.9% (*n* = 1 761) women have given the third birth
time_months, M = 33.03, SD = 14.70

Genders	Frequency	Percent (%)
Boy, boy	4 334	26.9
Girl, girl	3759	23.3
Boy, girl	4067	25.2
Girl, boy	3956	24.5

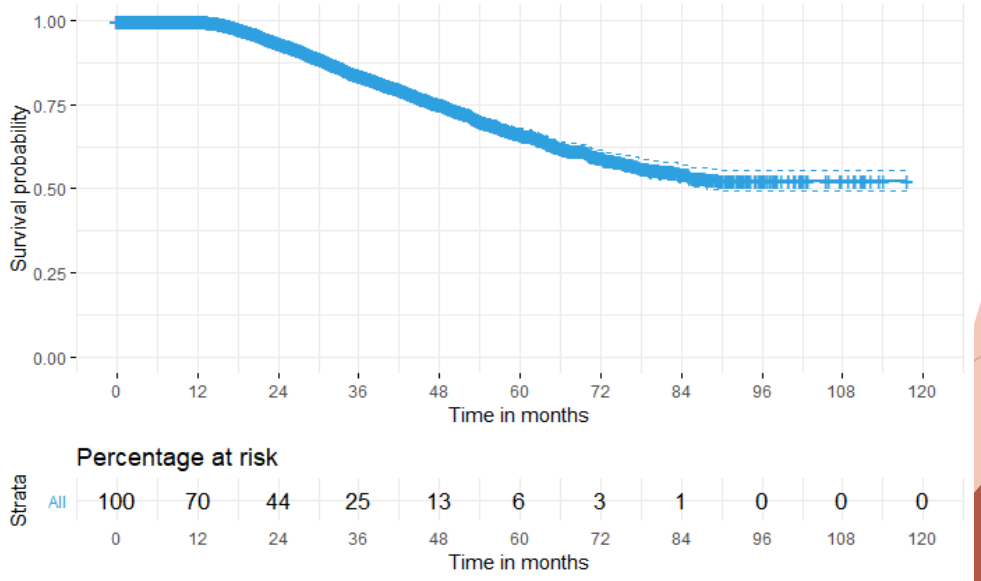
Survival analysis

What is the effect of gender of first two children on the likelihood of a woman having the third birth?

Objectives, methods, results

1. Comparing probability over time for having the 3rd birth after the 2nd birth later or not having it at all (*survival probability*) for combinations of the first two babies' genders.

Kaplan-Meier estimator plots:

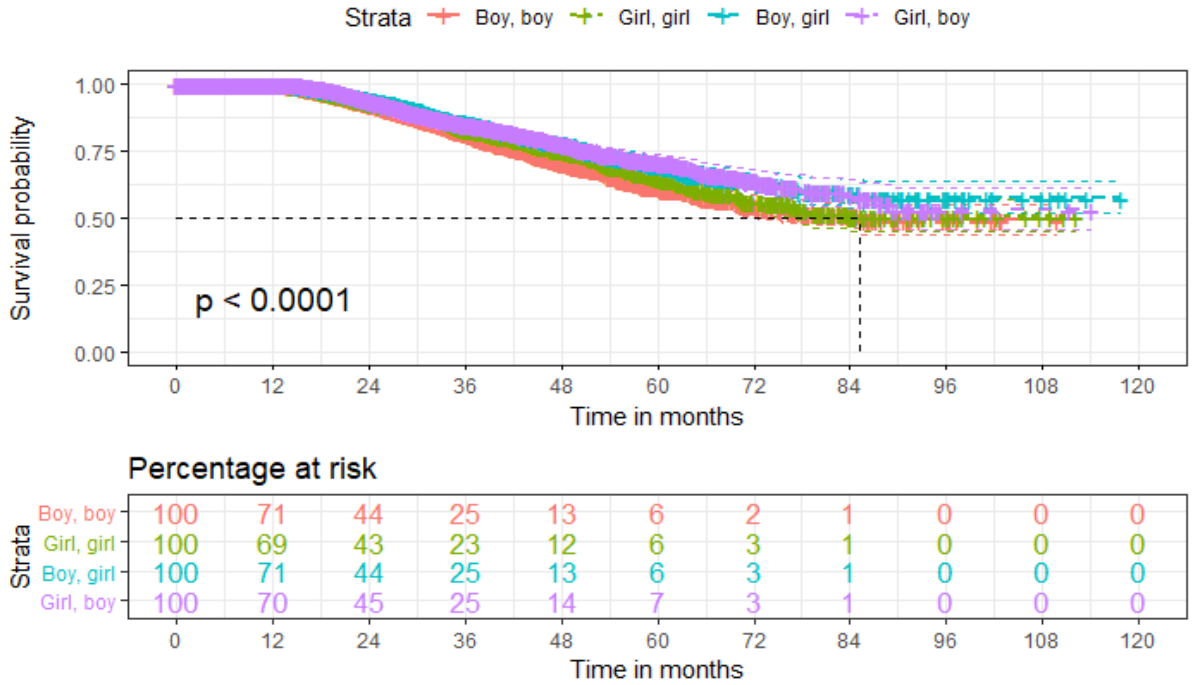


Survival analysis

What is the effect of gender of first two children on the likelihood of a woman having the third birth?

Objectives, methods, results

1. Comparing probability over time for having the 3rd birth after the 2nd birth later or not having it at all (*survival probability*) for combinations of the first two babies' genders.



Log-rank test:
χ^2 (3) = 26.4, p = 8e-06

		Weighted number of events (childbirth)	
Genders	n	Observed	Expected
Boy, boy	4 334	559	475
Girl, girl	3 759	411	397
Boy, girl	4 067	394	447
Girl, boy	3 956	397	442

Survival analysis

What is the effect of gender of first two children on the likelihood of a woman having the third birth?

Objectives, methods, results

2. Estimating a model for a survival time (from 2nd to 3rd birth), with respect to **age** of the mother at first birth, **time** between 1st and 2nd birth and **genders** of the first two children.

Multivariate Cox regression

Predictor	Regression coefficients					95.0% C.I. for Exp(B)	
	B	Exp(B)	Exp(-B)	S.E.(B)	Sig.	L.B.	U.B.
Age at first birth	.0049	1.0050	1.9951	.0145	.7341	.9768	1.0339
Time from 1st to 2nd birth	-.0003	.9997	1.0003	.0001	.0002	.9996	.9999
Boy, boy	.1174	1.1250	.8892	.06504	.0711	.9900	1.2774
Boy, girl	-.1671	.8461	1.1818	.0705	.0178	.7369	.9716
Girl, boy	-.1477	.8627	1.1592	.0704	.0359	.7515	.9903

N = 16 116, *n* of events = 1 761.

Abbreviations: B, regression coefficient; S.E., standard error; Sig., significance level; C.I., confidence interval; L.B., lower bound; U.B., upper bound.

Survival analysis

What is the effect of gender of first two children on the likelihood of a woman having the third birth?

Objectives, methods, results

3. Identifying the top ten women from the new unlabeled dataset of 500 who have the highest risk rates for having the third birth.

Rank	ID	Age at first birth	Time from 1st to 2nd birth (days)	Time from 1st to 2nd birth (months)	Gender	Risk score
1	308	23.3	381	12.49	Boy, boy	.328
2	130	24.0	415	13.61	Boy, boy	.323
3	386	19.8	357	11.70	Boy, boy	.317
4	412	28.0	515	16.89	Boy, boy	.316
5	370	19.2	361	11.84	Boy, boy	.313
6	298	22.4	429	14.07	Boy, boy	.311
7	205	26.6	507	16.62	Boy, boy	.311
8	93	27.1	526	17.25	Boy, boy	.308
9	37	20.7	411	13.48	Boy, boy	.307
10	405	26.4	527	17.28	Boy, boy	.305

Structural Equation Modelling

Post-Millennials' willingness to purchase products with geographical indications

Investigation on how

- culture,
- traditions,
- health benefits,
- perceived quality and safety, as well as
- support to rural development

affects **consumers' willingness to purchase hellim / halloumi cheese with GI**, and what is the mediating role of **trust in GI** in this relationship,

with specific interest in exploring these relationships among post-millennials.

Structural Equation Modelling

Post-Millennials' willingness to purchase products with geographical indications

Data: survey data with 238 rows

Construct	Item	Statement
Culture and traditions	CT1	Hellim with "Hellim/Halloumi" logo is a part of Cyprus culture
	CT2	For me it is important that hellim with "Hellim/Halloumi" logo is produced in traditional way
	CT3	Hellim with "Hellim/Halloumi" logo is made with traditional recipe
Health benefits	HB1	I think hellim with "Hellim/Halloumi" logo is higher quality
	HB2	I think hellim with "Hellim/Halloumi" logo is more natural
	HB3	I think hellim with "Hellim/Halloumi" logo is healthier
	HB4	I think hellim with "Hellim/Halloumi" logo is tastier
Rural development	RD1	By buying hellim with "Hellim/Halloumi" logo I support local community
	RD2	By buying hellim with "Hellim/Halloumi" logo I support local economy
	RD3	By buying hellim with "Hellim/Halloumi" logo I support local farmers
	RD4	By buying hellim with "Hellim/Halloumi" logo I contribute to preserve rural areas
Perceived quality and safety	PQS1	Hellim with "Hellim/Halloumi" logo is better quality than other hellim offered in the market
	PQS2	"Hellim/Halloumi" logo secures that I am buying quality product
	PQS3	Quality and safety of hellim is certified by "Hellim/Halloumi" logo
	PQS4	Quality and safety of hellim is guaranteed by "Hellim/Halloumi" logo
Trust	T1	I trust "Hellim/Halloumi" logo
	T2	"Hellim/Halloumi" logo is honest
	T3	"Hellim/Halloumi" logo is safe
Consumers' willingness to purchase	CWP1	I am willing to purchase hellim with "Hellim/Halloumi" logo
	CWP2	I am looking for hellim with "Hellim/Halloumi" logo in the markets

Caputo et al, 2011; Teuber, 2011; Mazodier, 2015; Lai, & Cheng, 2016

Structural Equation Modelling

Post-Millennials' willingness to purchase products with geographical indications

Reliability, and convergent and discriminant validity of the constructs

Subscale	Cronbach's α
Culture and traditions	.50
Health benefits	.95
Rural development	.90
Perceived quality and safety	.93
Trust	.95
Consumer's willingness to purchase	.80

	CR	AVE	MSV	MaxR(H)	PQS	CT	HB	RD	CWP	T
PQS	0.929	0.768	0.590	0.948	**0.876**					
CT	0.492	0.254	0.449	0.524	0.623	**0.504**				
HB	0.950	0.827	0.551	0.955	0.742	0.655	**0.909**			
RD	0.913	0.681	0.410	0.943	0.640	0.568	0.486	**0.825**		
CWP	0.812	0.684	0.503	0.820	0.633	0.545	0.625	0.509	**0.827**	
T	0.950	0.865	0.590	0.954	0.768	0.670	0.666	0.565	0.709	**0.930**

CR = Composite Reliability;
AVE = Average Variance Extracted;
MSV = Maximum Shared Variance;
MaxR(H) = McDonald Construct Reliability;
SQRT(AVE) on diagonal.

VALIDITY CONCERNS

Discriminant Validity: the square root of the AVE for CT is less than one the absolute value of the correlations with another factor.

Reliability: the CR for CT is less than 0.70.

Convergent Validity: the AVE for CT is less than 0.50.

Discriminant Validity: the AVE for CT is less than the MSV.

Structural Equation Modelling

Post-Millennials' willingness to purchase products with geographical indications

Adequate measurement and structural model fit

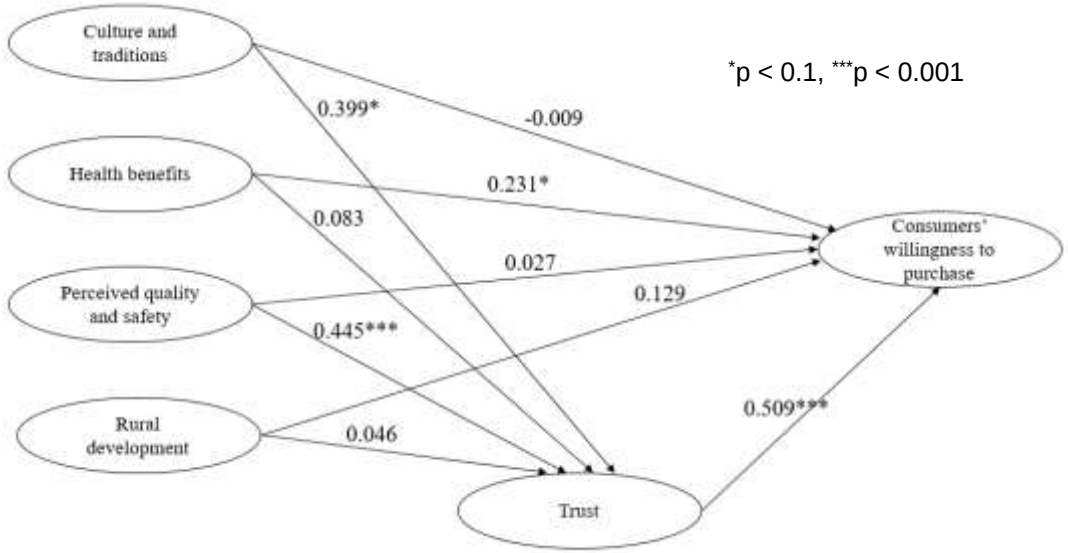
	Measurement model fit (RD5 excluded)	Structural model fit
χ2/df	1.948	1.948
RMSEA	.063	.063
SRMR	.042	.042
AGFI	.844	.849
NFI, TLI, CFI	> .90	> .90

RMSEA, root mean square error of approximation;
SRMR, standardized root mean residual;
AGFI, adjusted goodness of fit index;
NFI, Bentler-Bonett Normed Fit Index;
TLI, Tucker-Lewis index;
CFI, comparative fit index

Structural Equation Modelling

Post-Millennials' willingness to purchase products with geographical indications

Structural equation model



*p < 0.1, ***p < 0.001

Mediation

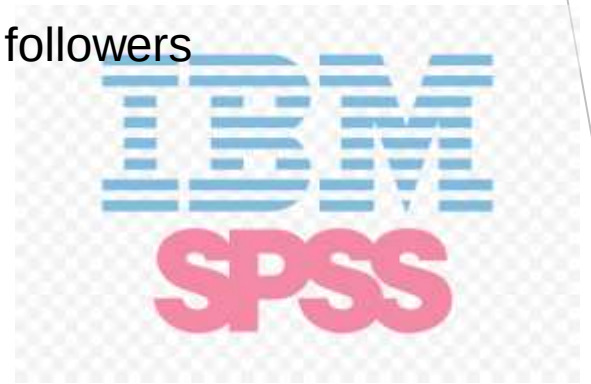
1. Baron, & Kenny, 1986 (first two columns)
2. the bias-corrected bootstrap method with 10.000 samples at 95% confidence intervals (third column)

Relationship	Standardized regression weights without mediator	Standardized regression weights with mediator	Standardized indirect effects
CT -> T -> CWP	0.160	-0.006	0.012**
HB -> T -> CWP	0.222**	0.246**	-0.074
PQS -> T -> CWP	0.299***	0.028	0.072***
RD -> T -> CWP	0.110	0.100	-0.067

**p < 0.1
***p < 0.01

Survey experiment data analysis

Effects of the type of Instagram endorser on the purchasing intentions of their followers

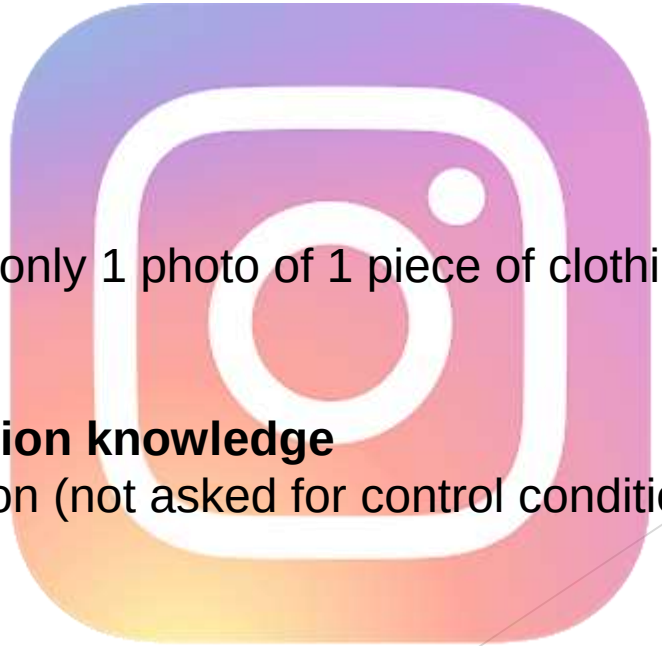


Data: experimental survey data with 203 rows

Survey
4 groups:

1. Celebrity
2. Influencer
3. Peer
4. Control condition

4 photos of clothing, but each respondent saw only 1 photo of 1 piece of clothing



- Asked about their **purchase intention**
- Asked if they feel being influenced = **persuasion knowledge**
- Asked about their **perception** about the person (not asked for control condition)
- Asked about they are **fashion-conscious**
- Asked about their Instagram **activity**
- Asked about **sex** and **age**

Survey experiment data analysis

Effects of the type of Instagram endorser on the purchasing intentions of their followers

Data: experimental survey data with 203 rows

Survey
4 groups:

1. Celebrity
2. Influencer
3. Peer
4. Control condition

Survey experiment data analysis

Effects of the type of Instagram endorser on the purchasing intentions of their followers

Data: experimental survey data with 203 rows

Survey
4 groups:

1. Celebrity
2. Influencer
3. Peer
4. Control condition

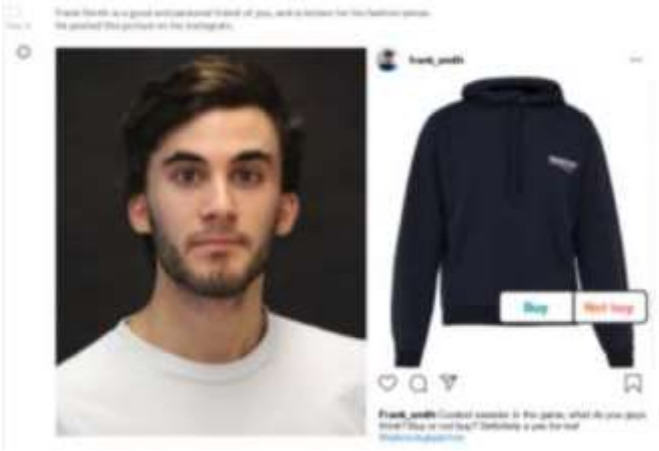
Survey experiment data analysis

Effects of the type of Instagram endorser on the purchasing intentions of their followers

Data: experimental survey data with 203 rows

Survey
4 groups:

1. Celebrity
2. Influencer
3. Peer
4. Control condition



Survey experiment data analysis

Effects of the type of Instagram endorser on the purchasing intentions of their followers

Data: experimental survey data with 203 rows

Survey
4 groups:

1. Celebrity
2. Influencer
3. Peer
4. Control condition

Survey experiment data analysis

Effects of the type of Instagram endorser on the purchasing intentions of their followers

Data: experimental survey data with 203 rows

Survey

4 groups:

1. Celebrity
2. Influencer
3. Peer
4. Control condition

4 photos of clothing, but each respondent saw only 1 photo of 1 piece of clothing

- Asked about their **purchase intention**
- Asked if they feel being influenced = **persuasion knowledge**
- Asked about their **perception** about the person (not asked for control condition)
- Asked about they are **fashion-conscious**
- Asked about their Instagram **activity**
- Asked about **sex** and **age**

Survey experiment data analysis

Effects of the type of Instagram endorser on the purchasing intentions of their followers

Reliability of the scales (internal consistency)

- Purchase intention (2), ρ = .91
- Perception (6), α = .70
  - Attractiveness (2), ρ = .72
  - Expertise (2), ρ = .23
  - Reliability (2), ρ = .89
- Fashion consciousness (3), α = .82

ρ, Spearman-Brown split-half coefficient
α, Cronbach's coefficient

Survey experiment data analysis

Effects of the type of Instagram endorser on the purchasing intentions of their followers

Equality of experimental conditions by gender and age

	Gender								
	Female			*Male*			Total		
Experimental condition	Mean	SD	N	Mean	SD	N	Mean	SD	N
Control	23.82	2.519	22	26.00	3.212	26	25.00	3.087	48
Celebrity	24.33	3.464	27	25.87	3.005	23	25.04	3.319	50
Influencer	23.53	2.423	32	26.33	3.464	18	24.54	3.118	50
Peer	23.93	5.298	28	24.53	3.289	19	24.17	4.560	47
Total	23.89	3.594	109	25.71	3.239	86	24.69	3.550	195

SD = standard deviation.

Chi-squared test of independence:
no association between ***gender*** and ***experimental condition***, χ^2 (3) = 3.63, p = .30

ANOVA of age:
effect of ***gender***, F (1, 187) = 12.45, p = .001, partial η^2 = .06
no effect of ***group***, F (3, 187) = .58, p = .63, partial η^2 = .01
no effect of ***gender × group***, F (3, 187) = .85, p = .47, partial η^2 = .01

Survey experiment data analysis

Effects of the type of Instagram endorser on the purchasing intentions of their followers

Effects of the type of endorser on the purchase intention – ANOVA

Assumptions of normality and absence of outliers satisfied.

No between-subjects effect of ***group*** on ***purchase intention***,

F (3, 195) = .68, p = .57.

Experimental condition	Mean	SD	SE	95% Confidence Interval for Mean	
				Lower Bound	Upper Bound
Control	3.19	1.133	.160	2.87	3.51
Celebrity	3.37	1.138	.161	3.05	3.69
Influencer	3.45	1.097	.154	3.14	3.76
Peer	3.48	1.096	.158	3.16	3.80
Total	3.37	1.113	.079	3.22	3.53

SD, standard deviation. SE, standard error.

Survey experiment data analysis

Effects of the type of Instagram endorser on the purchasing intentions of their followers

Effects of the type of endorser on the persuasion knowledge (ordinal variable) – Kruskal-Wallis test

Significant effect of ***group*** on ***persuasion knowledge***,

χ^2 (3) = 73.20, p < .001.

Experimental condition	Mean	SD	Mean rank
Control	1.62	.901	61.17
Celebrity	2.90	1.093	121.91
Influencer	3.45	1.137	142.35
Peer	1.83	.953	72.63
Total	2.46	1.270	

Mann-Whitney U test:

non-significant difference only between ’control’ and ’peer’ group (p > .05)

Survey experiment data analysis

Effects of the type of Instagram endorser on the purchasing intentions of their followers

Effects of the type of endorser on the perception – MANOVA (***attractiveness***, ***expertise*** and ***reliability*** as DVs)

Assumptions of sample size, univariate normality, multivariate normality, linear relationships among DVs, multicollinearity/singularity, equality of covariance matrices (Box's test) – not importantly violated.

Assumption of equality of error variances not violated only for ***reliability*** (Levene: $p > .05$).

Solution: criterion for between-subjects effects tests was set at $p = .01$.

Survey experiment data analysis

Effects of the type of Instagram endorser on the purchasing intentions of their followers

Effects of the type of endorser on the perception – MANOVA (***attractiveness***, ***expertise*** and ***reliability*** as DVs)

	Perception dimension					
	Attractiveness		*Expertise*		*Reliability*	
Experimental condition	Mean	SD	Mean	SD	Mean	SD
Celebrity	3.95	.970	3.26	.777	2.93	.742
Influencer	3.75	.751	3.25	1.036	3.00	.728
Peer	3.55	.573	3.33	.574	4.21	.764
Total	3.76	.796	3.28	.819	3.36	.942

Multivariate tests:

significant effect of ***group***, Pillai's Trace = .51, F (6, 288) = 16.35, $p < .001$, partial $\eta^2 = .25$

Tests of between-subjects effects:

- no effect of *group* on ***attractiveness***, F (2, 145) = 3.09, $p = .05$, partial $\eta^2 = .04$
- no effect of group on ***expertise***, F (2, 145) = .14, $p = .86$, partial $\eta^2 = .002$
- effect of *group* on ***reliability***, F (2, 145) = 45.13, $p < .001$, partial $\eta^2 = .38$ (peer highest)

Survey experiment data analysis

Effects of the type of Instagram endorser on the purchasing intentions of their followers

Effects of persuasion knowledge, attractiveness, expertise, reliability, fashion consciousness, activity and gender on the purchase intention, when the type of endorser is controlled – multiple regression

Assumptions of sample size, multicollinearity/singularity, absence of atypical points, normality, linearity, homogeneity of variances, and independence of residuals were not violated.

IVs explained 61.3% of the variance of ***purchase intention***,

$R^2 = .61$, $F\ (10, 136) = 21.51$, $p < .001$.

Survey experiment data analysis

Effects of the type of Instagram endorser on the purchasing intentions of their followers

Effects of persuasion knowledge, attractiveness, expertise, reliability, fashion consciousness, activity and gender on the purchase intention, when the type of endorser is controlled – multiple regression

	Unstandardized Coefficients		Std. Coefficients			95.0% C.I. for B	
	B	Std. Error	Beta	t	p	L.B.	U.B.
(Constant)	-.798	.491		-1.624	.107	-1.770	.174
1. Celebrity	.327	.191	.127	1.713	.089	-.051	.704
2. Influencer	.311	.204	.122	1.528	.129	-.092	.715
3. Peer	-.069	.209	-.027	-.332	.741	-.483	.345
4. Gender	-.157	.134	-.070	-1.173	.243	-.421	.108
5. Persuasion knowledge	-.038	.061	-.043	-.612	.541	-.159	.084
6. Attractiveneess	.018	.095	.013	.190	.850	-.170	.206
7. Expertise	.293	.089	.216	3.298	.001	.117	.469
8. Reliability	.228	.108	.193	2.108	.037	.014	.441
9. Fashion consciousness	.680	.097	.515	7.011	.000	.488	.872
10. Activity	-.024	.061	-.025	-.396	.693	-.145	.097

C.I., confidence interval; L.B., lower bound; U.B., upper bound.

Survey experiment data analysis

Effects of the type of Instagram endorser on the purchasing intentions of their followers

Effects of persuasion knowledge, attractiveness, expertise, reliability, fashion consciousness, activity and gender on the purchase intention, when the type of endorser is controlled – multiple regression

	Unstandardized Coefficients		Std. Coefficients			95.0% C.I. for B	
	B	Std. Error	Beta	t	p	L.B.	U.B.
(Constant)	-.798	.491		-1.624	.107	-1.770	.174
1. Celebrity	.327	.191	.127	1.713	.089	-.051	.704
2. Influencer	.311	.204	.122	1.528	.129	-.092	.715
3. Peer	-.069	.209	-.027	-.332	.741	-.483	.345
4. Gender	-.157	.134	-.070	-1.173	.243	-.421	.108
5. Persuasion knowledge	-.038	.061	-.043	-.612	.541	-.159	.084
6. Attractiveneess	.018	.095	.013	.190	.850	-.170	.206
7. Expertise	.293	.089	.216	3.298	.001	.117	.469
8. Reliability	.228	.108	.193	2.108	.037	.014	.441
9. Fashion consciousness	.680	.097	.515	7.011	.000	.488	.872
10. Activity	-.024	.061	-.025	-.396	.693	-.145	.097

C.I., confidence interval; L.B., lower bound; U.B., upper bound.

Tableau Story

Profiles of the UK bank customers in different regions

Data: data for an imaginary bank operating in the UK, 4 014 rows

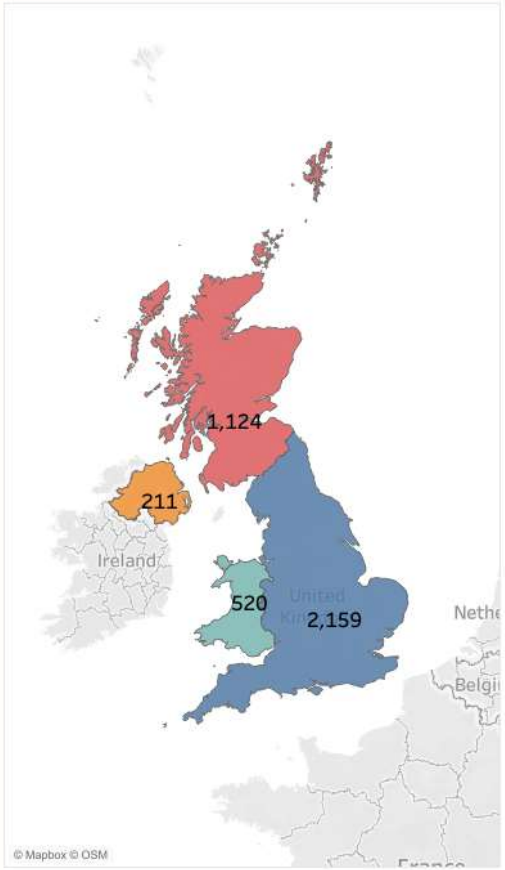
https://www.kaggle.com/ukveteran/uk-bank-customers

	A	B	C	D	E	F	G	H	I
1	Customer ID	Name	Surname	Gender	Age	Region	Job Classification	Date Joined	Balance
2	100000001	Simon	Walsh	Male	21	England	White Collar	05.Jan.15	113810.2
3	400000002	Jasmine	Miller	Female	34	Northern	Blue Collar	06.Jan.15	36919.73
4	100000003	Liam	Brown	Male	46	England	White Collar	07.Jan.15	101536.8
5	300000004	Trevor	Parr	Male	32	Wales	White Collar	08.Jan.15	1421.52
6	100000005	Deirdre	Pullman	Female	38	England	Blue Collar	09.Jan.15	35639.79
7	300000006	Ava	Coleman	Female	30	Wales	Blue Collar	09.Jan.15	122443.8
8	100000007	Dorothy	Thomson	Female	34	England	Blue Collar	11.Jan.15	42879.84
9	200000008	Lisa	Knox	Female	48	Scotland	Other	11.Jan.15	36680.17
10	300000009	Ruth	Campbell	Female	33	Wales	White Collar	11.Jan.15	74284.35

Question: How the customers look like in the different regions of the UK?

UK bank customer profiles by region

Distribution of the customers by region	General UK customer profile	Customers in England are somewhat younger than general, mostly white collars	Customers in Scotland are mostly men in their late 40s and 50s, mostly blue collars and others



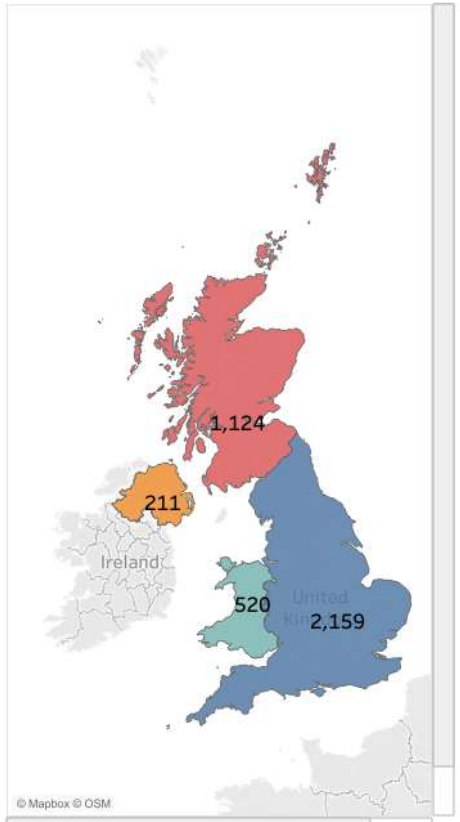
Region
- England
- Northern Ireland
- Scotland
- Wales

tableau

UK bank customer profiles by region

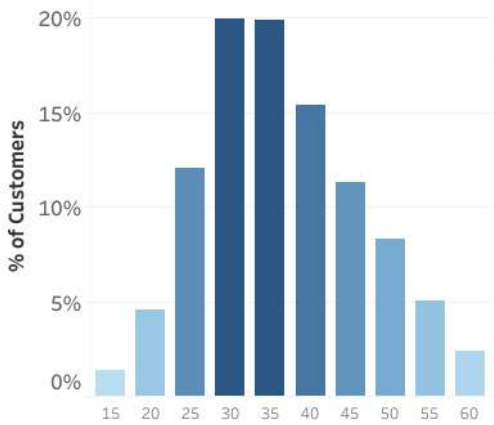
Distribution of the customers by region | General UK customer profile | Customers in England are somewhat younger than general, mostly white collars | Customers in Scotland are mostly men in their late 40s and 50s, mostly blue collars and others

Map

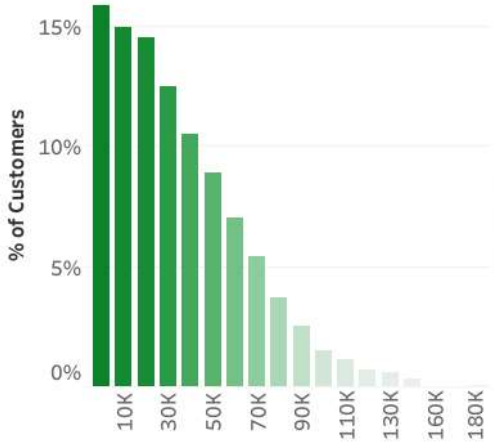


England
Wales
Northern Ireland
Scotland

Distribution by age



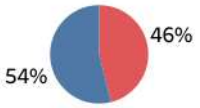
Distribution by balance



Age Groups
5

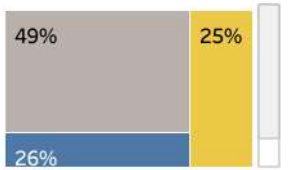
Balance Groups
10,000

Gender



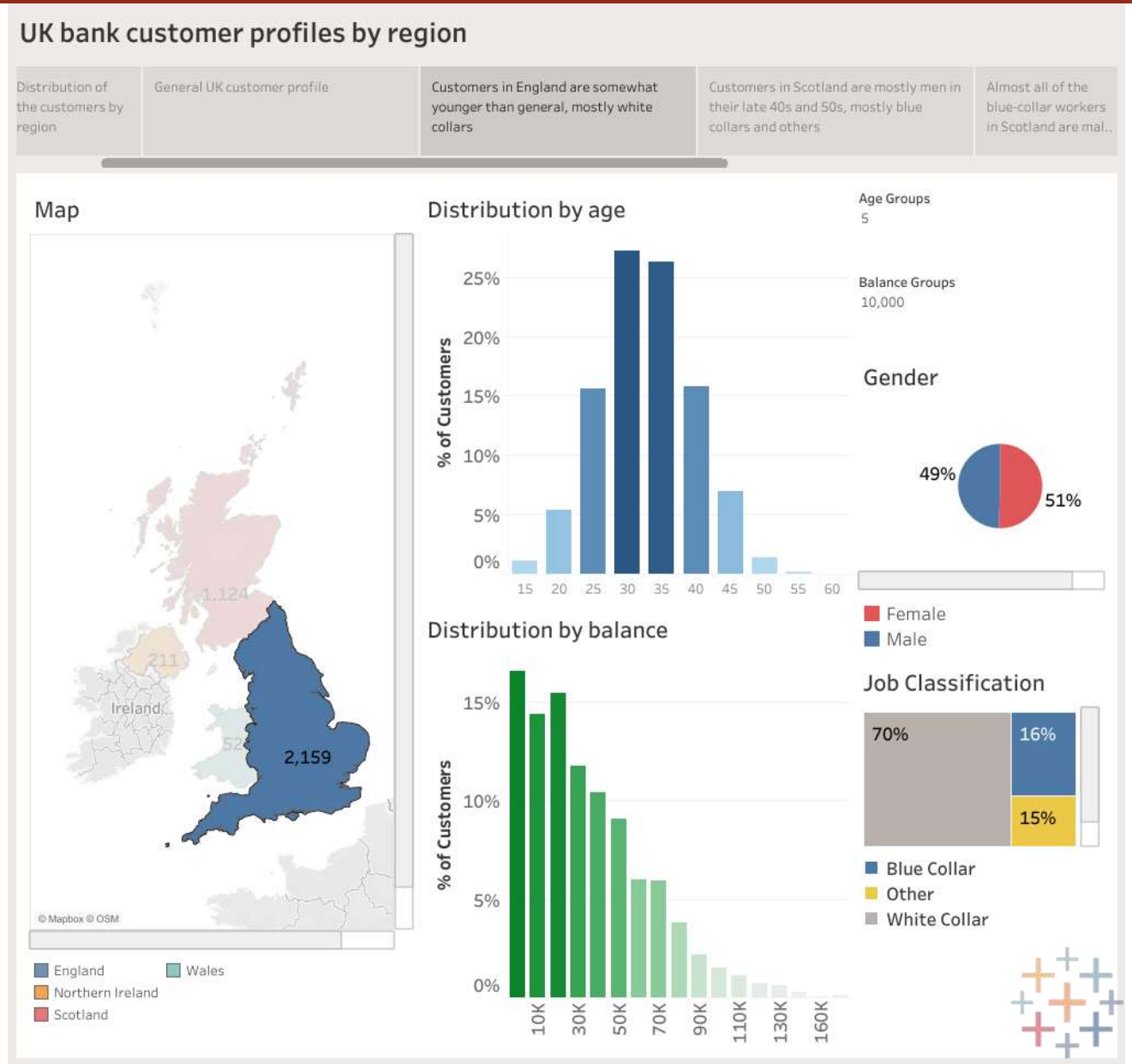
Female
Male

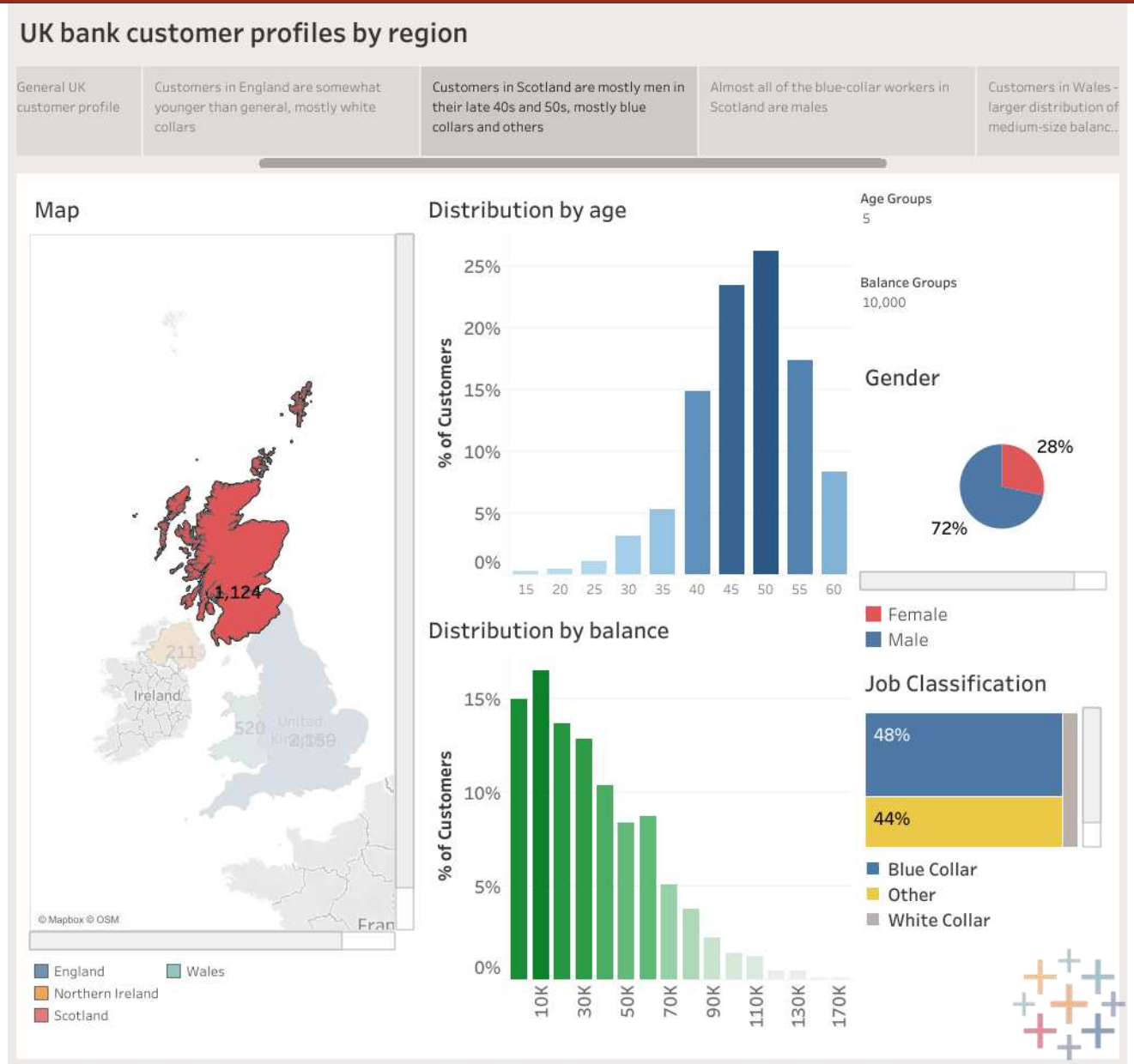
Job Classification

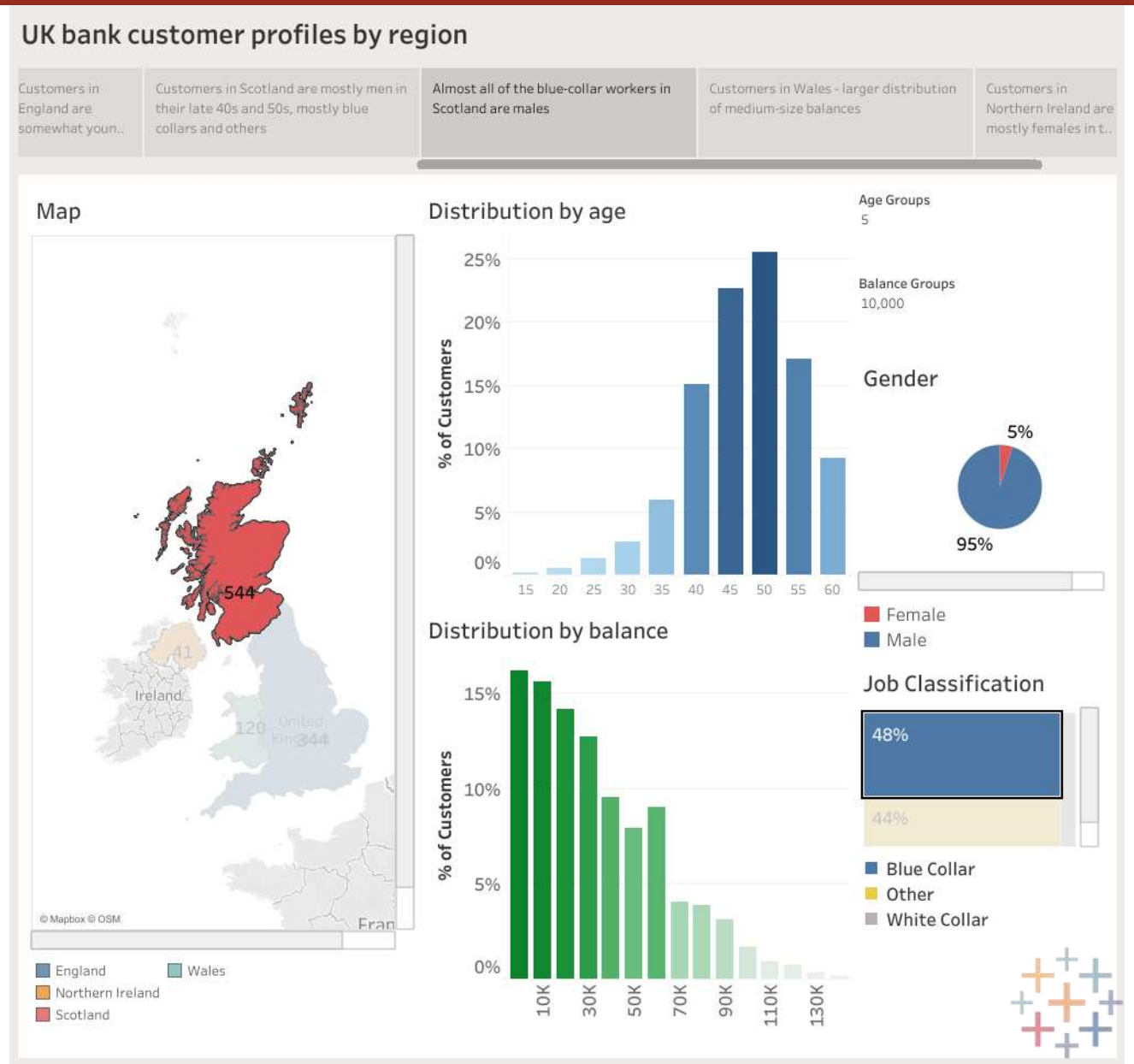


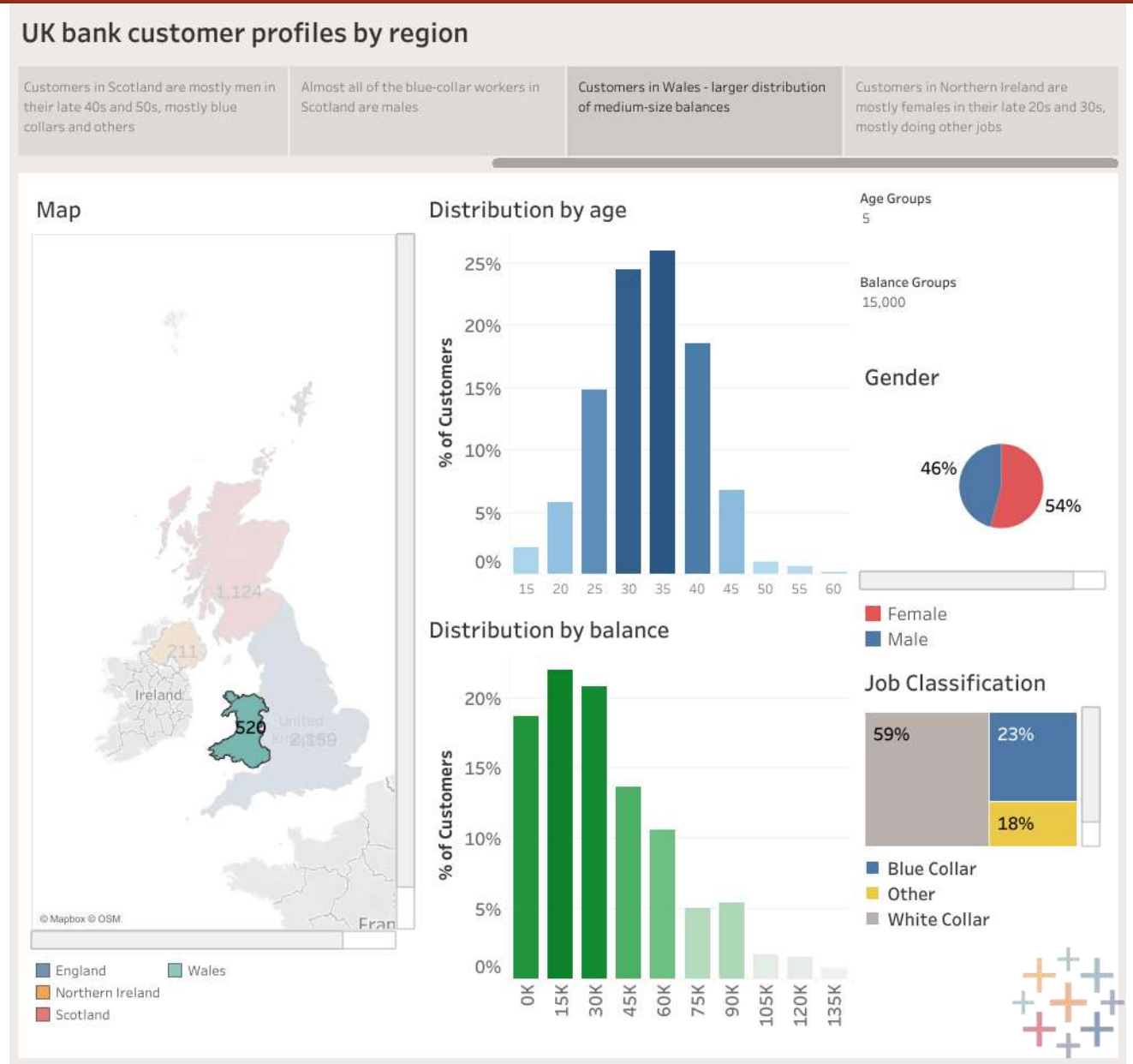
Blue Collar
Other
White Collar

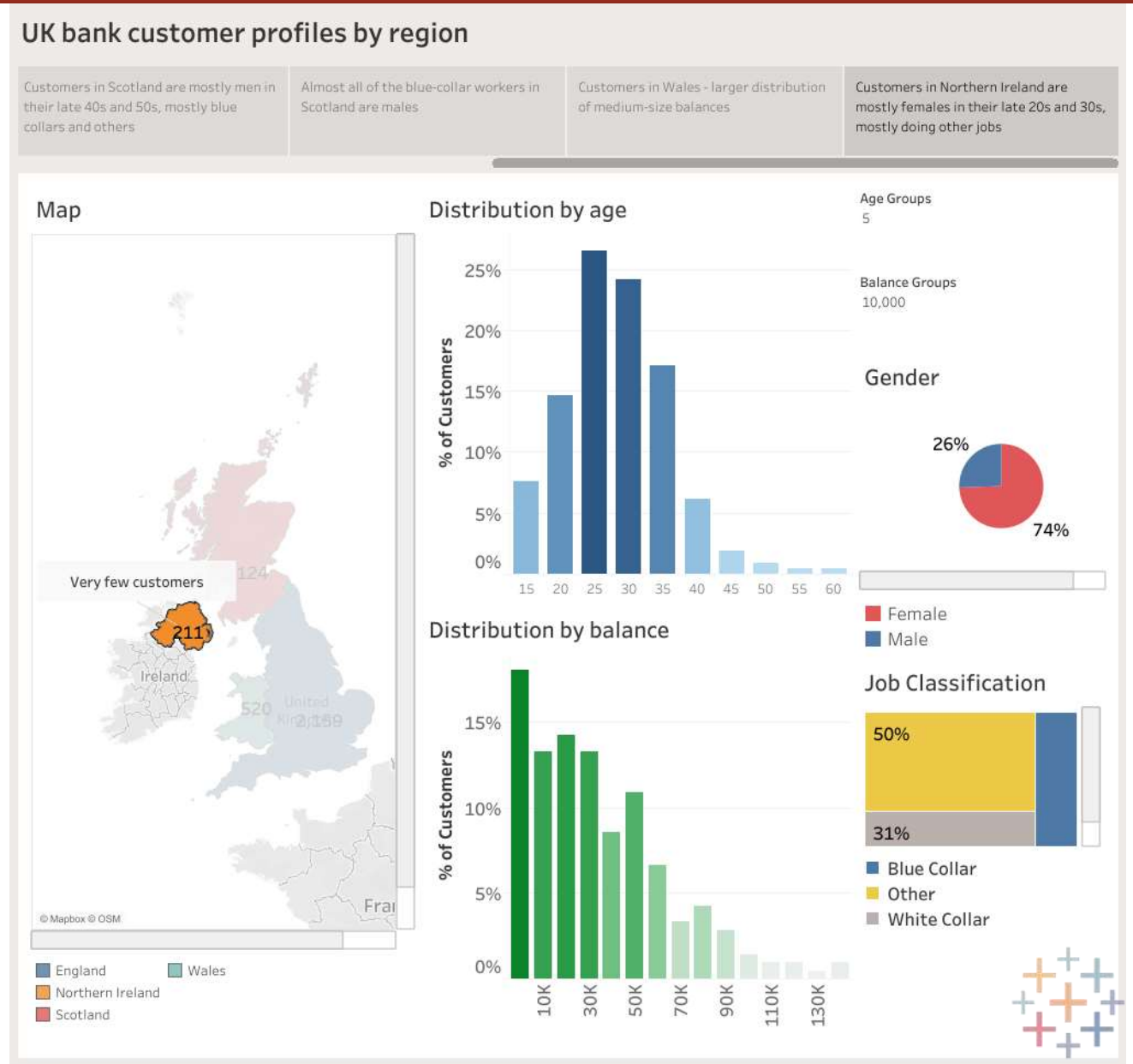
tableau


UK bank customer profiles by region
Distribution of the customers by region
General UK customer profile
Customers in England are somewhat younger than general, mostly white collars
Customers in Scotland are mostly men in their late 40s and 50s, mostly blue collars and others
Almost all of the blue-collar workers in Scotland are mal..
Map
1,124
211
Ireland
52
2,159
© Mapbox © OSM
England
Wales
Northern Ireland
Scotland
Distribution by age
% of Customers
25%
20%
15%
10%
5%
0%
15 20 25 30 35 40 45 50 55 60
Distribution by balance
% of Customers
15%
10%
5%
0%
10K 30K 50K 70K 90K 110K 130K 160K
Age Groups
5
Balance Groups
10,000
Gender
49%
51%
Female
Male
Job Classification
70%
16%
15%
Blue Collar
Other
White Collar
+ a b | e a u


UK bank customer profiles by region
General UK customer profile
Customers in England are somewhat younger than general, mostly white collars
Customers in Scotland are mostly men in their late 40s and 50s, mostly blue collars and others
Almost all of the blue-collar workers in Scotland are males
Customers in Wales - larger distribution of medium-size balanc...
Map
1,124
211
Ireland
520
2,159
© Mapbox © OSM
England
Wales
Northern Ireland
Scotland
Distribution by age
% of Customers
25%
20%
15%
10%
5%
0%
15 20 25 30 35 40 45 50 55 60
Distribution by balance
% of Customers
15%
10%
5%
0%
10K 30K 50K 70K 90K 110K 130K 170K
Age Groups
5
Balance Groups
10,000
Gender
28%
72%
Female
Male
Job Classification
48%
44%
Blue Collar
Other
White Collar


UK bank customer profiles by region
Customers in England are somewhat youn..
Customers in Scotland are mostly men in their late 40s and 50s, mostly blue collars and others
Almost all of the blue-collar workers in Scotland are males
Customers in Wales - larger distribution of medium-size balances
Customers in Northern Ireland are mostly females in t..
Map
544
41
120
344
Ireland
© Mapbox © OSM
England
Wales
Northern Ireland
Scotland
Distribution by age
% of Customers
25%
20%
15%
10%
5%
0%
15 20 25 30 35 40 45 50 55 60
Distribution by balance
% of Customers
15%
10%
5%
0%
10K 30K 50K 70K 90K 110K 130K
Age Groups
5
Balance Groups
10,000
Gender
5%
95%
Female
Male
Job Classification
48%
44%
Blue Collar
Other
White Collar
tableau


UK bank customer profiles by region
Customers in Scotland are mostly men in their late 40s and 50s, mostly blue collars and others
Almost all of the blue-collar workers in Scotland are males
Customers in Wales - larger distribution of medium-size balances
Customers in Northern Ireland are mostly females in their late 20s and 30s, mostly doing other jobs
Map
1,124
211
Ireland
520
2,159
© Mapbox © OSM
England
Wales
Northern Ireland
Scotland
Distribution by age
% of Customers
25%
20%
15%
10%
5%
0%
15 20 25 30 35 40 45 50 55 60
Distribution by balance
% of Customers
20%
15%
10%
5%
0%
0K 15K 30K 45K 60K 75K 90K 105K 120K 135K
Age Groups
5
Balance Groups
15,000
Gender
46%
54%
Female
Male
Job Classification
59%
23%
18%
Blue Collar
Other
White Collar


UK bank customer profiles by region
Customers in Scotland are mostly men in their late 40s and 50s, mostly blue collars and others
Almost all of the blue-collar workers in Scotland are males
Customers in Wales - larger distribution of medium-size balances
Customers in Northern Ireland are mostly females in their late 20s and 30s, mostly doing other jobs
Map
Very few customers
211
124
520
Ireland
© Mapbox © OSM
England
Wales
Northern Ireland
Scotland
Distribution by age
% of Customers
25%
20%
15%
10%
5%
0%
15 20 25 30 35 40 45 50 55 60
Distribution by balance
% of Customers
15%
10%
5%
0%
10K 30K 50K 70K 90K 110K 130K
Age Groups
5
Balance Groups
10,000
Gender
26%
74%
Female
Male
Job Classification
50%
31%
Blue Collar
Other
White Collar

Thank you!