

Pig E. Bank Client Loss Analysis Report

Objective

Identify top factors that indicate client loss at Pig E. Bank to increase customer retention.

Data Quality Assessment

Data Dimensions Before Cleaning: 14 columns, 991 rows

Documentation of Cleaning:

Columns	Description	Mean	Max	Min	Missing Values	Quality Check	Missing Values	Duplicates
Row_Number	A numerical identifier for the rows					This column was unnecessary, therefore deleted.		No duplicates were found in the data set.
Customer_ID	Unique identifier for customers				0	Renamed: Customer ID	retained without	
Last_Name	Last name of the customer				1	Renamed: Last Name.		
Credit Score	Numerical score representing the creditworthiness of the customer	648,512,146	850	376	3		Imputed missing credit scores with the mean	
Country	Customer's residential country				0	Country abbreviations were replaced with their corresponding standard full		
Gender	Gender of the customer				1	Gender abbreviations ('M' and 'F') were replaced with the full terms 'Male' and		
Age	Age of the customer	386,848,485	82	2	2	recorded as '2' that need	and 2-year-old with	
Tenure	Number of years the customer has been with the bank	506,659,939	10	0	0			
Balance	Current balance in the customer's account	780,027,181	213146,2	0	0			
NumOfProducts	Number of products the customer has	152,270,434	4	1	0	Rename column from 'NumOfProducts' to 'Number of Products'		
HasCrCard?	Indicates whether the customer has a credit card (1) or not (0)				0	Rename column from 'HasCrCard?' to 'Has Credit		
IsActiveMember	Indicates whether the customer has been active (1) or not (0)				0	'IsActiveMember' to 'Is		
Estimated Salary	Estimated salary of the customer	98,574,543	199725,39	371,05	0		Imputed missing estimated salaries the	
ExitedFromBank?	Indicates whether the customer has left the bank (1) or not (0)				0	Rename column from		

Columns Added	Description
Credit Score Group	This column categorizes credit scores into ranges of 99 points each, such as 600-699 and 700-799, to simplify analysis and identify trends more easily.
Age Group	This column categorizes ages into groups, with most spanning 9 years, such as 25-34 and 35-44, except for the initial group of 18-24, to facilitate demographic analysis and trend identification.
Balance Group	This column categorizes balances into groups of \$10,000 increments, such as 0-10,000, 10,001-20,000, to streamline financial analysis and comparison.
Salary Group	This column categorizes estimated salaries into into groups of \$10,000 increments, such as 0-10,000, 10,001-20,000, to streamline financial analysis and comparison.

note: no duplicates found the Pig E. Bank client data set.

Descriptive Statistics

Stayed At Bank (774 Total): Current customers

Column	Min	Max	Average	Median	Mode
Credit Score	411	850	652	657	850
Age	18	82	38	36	36
Tenure	0	10	5	5	9
Balance	\$0.00	\$197,041.80	\$75,054.79	\$93,708.58	\$0.00
NumOfProducts	1	3	2	2	2
Estimated Salary	\$371.05	\$199,661.50	\$99,472.72	\$98,733.31	N/A

Exited From Bank (202 Total): Former customers

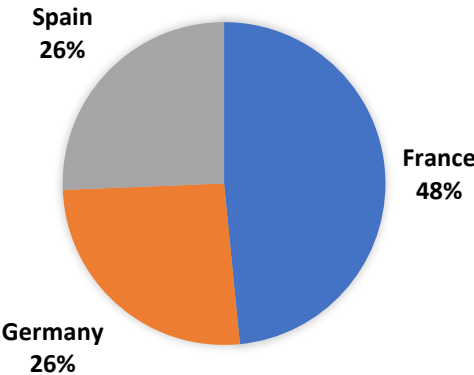
Column	Min	Max	Average	Median	Mode
Credit Score	376	850	637	644	750
Age	22	69	45	45	39
Tenure	0	10	5	5	1
Balance	\$0.00	\$213,146.20	\$90,703.94	\$112,928.24	\$0.00
NumOfProducts	1	4	1	1	1
Estimated Salary	\$417.41	\$199,725.39	\$97,446.03	\$100,375.40	N/A

Exploring Variables: Demographics

Country:

		All Customers	Current Customers	Former Customers
Country	France	48.44%	51.21%	37.75%
	Germany	25.93%	23.13%	36.76%
	Spain	25.63%	25.67%	25.49%

TOTAL CUSTOMER BY COUNTRY



Out of the customers that left, most of them were from France. This makes sense because France also has the largest customer base.

37.7% of customers from France have left.

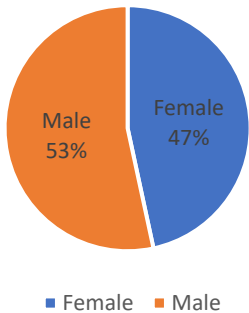
Germany, however, has a similar number of customers leaving, but a smaller customer base.

36.7% of customers from Germany have left.

Gender:

Gender	All Customers	Current Customers	Former Customers
Female	46.62%	43.33%	59.31%
Male	53.28%	56.54%	40.69%
Unknown	0.10%	0.13%	

Total customer by gender

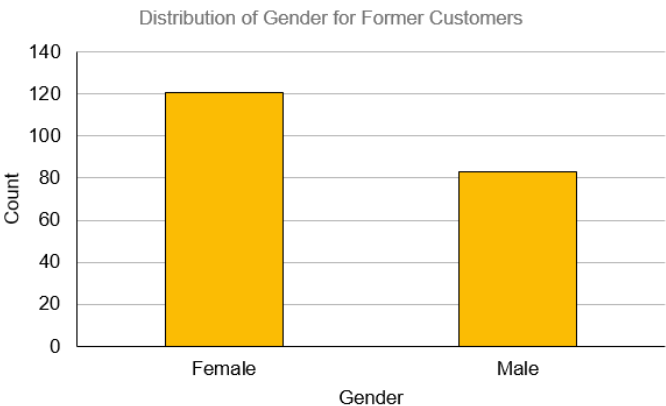
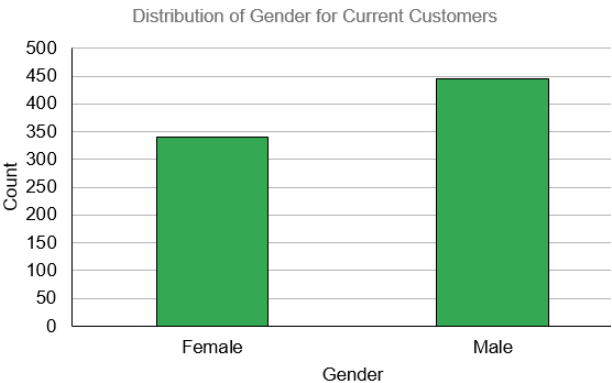
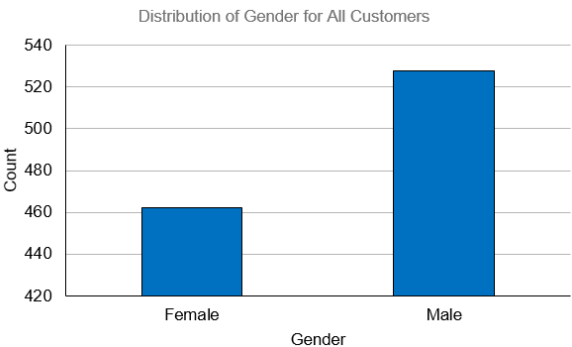
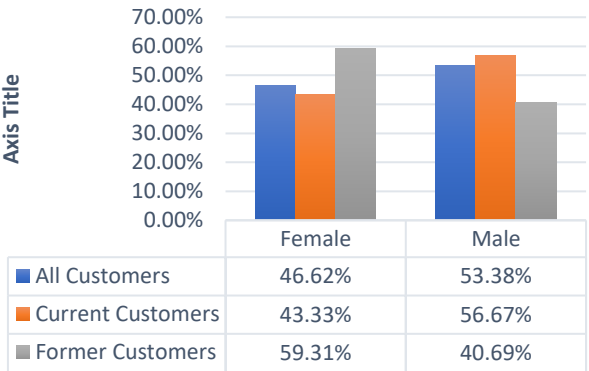


There were more female customers that left than male customers.

Of all female customers, 59.31% left. Of all male customers, 40.69% left.

Female customers also make up less of the customer base than male customers.

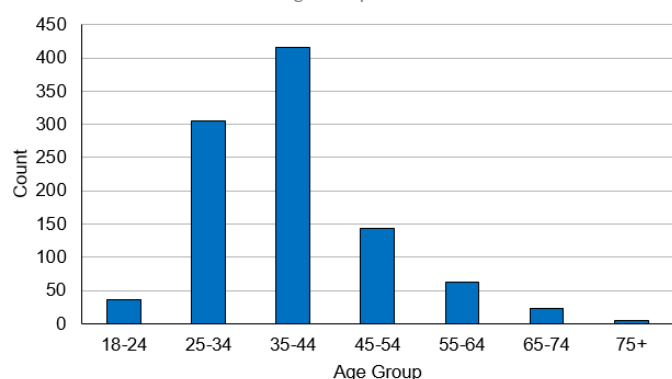
Customer by Gender



Age:

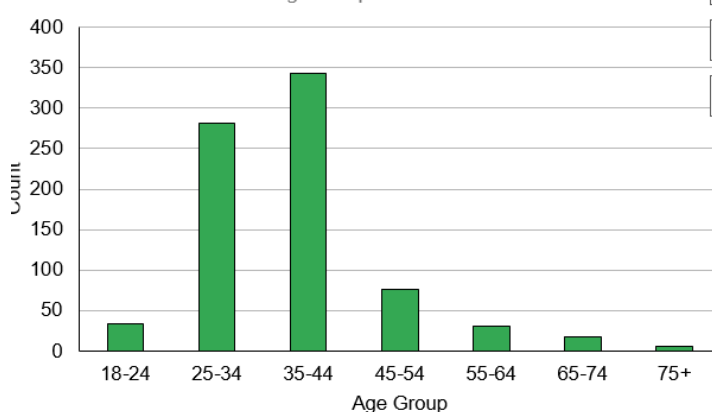
	All Customers	Current Customers	Former Customers
18-24	3.63%	4.32%	0.98%
25-34	30.78%	35.71%	11.76%
35-44	41.98%	43.58%	35.78%
45-54	14.53%	9.66%	33.33%
55-64	6.36%	3.94%	15.69%
65-74	2.22%	2.16%	2.45%
75+	0.50%	0.64%	

Distribution of Age Group for All Customers



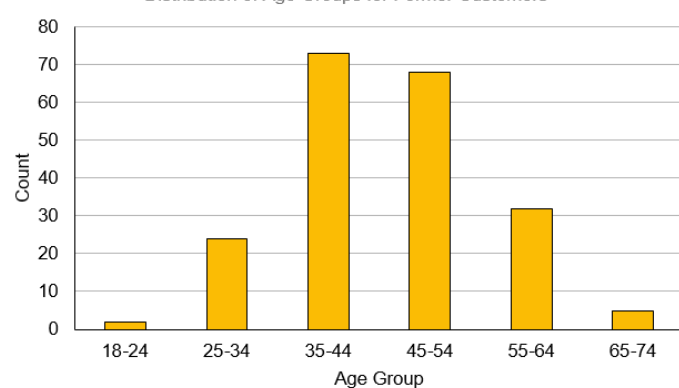
Approximately 80% of current customers are aged 25-44, while only about 45% of former customers fall within this age range. Customers aged 45-64 constitute about 13% of current customers but represent approximately 58% of former customers. Furthermore, customers between 35-54 years old account for around 70% of those who have left the bank.

Distribution of Age Group for Current Customers



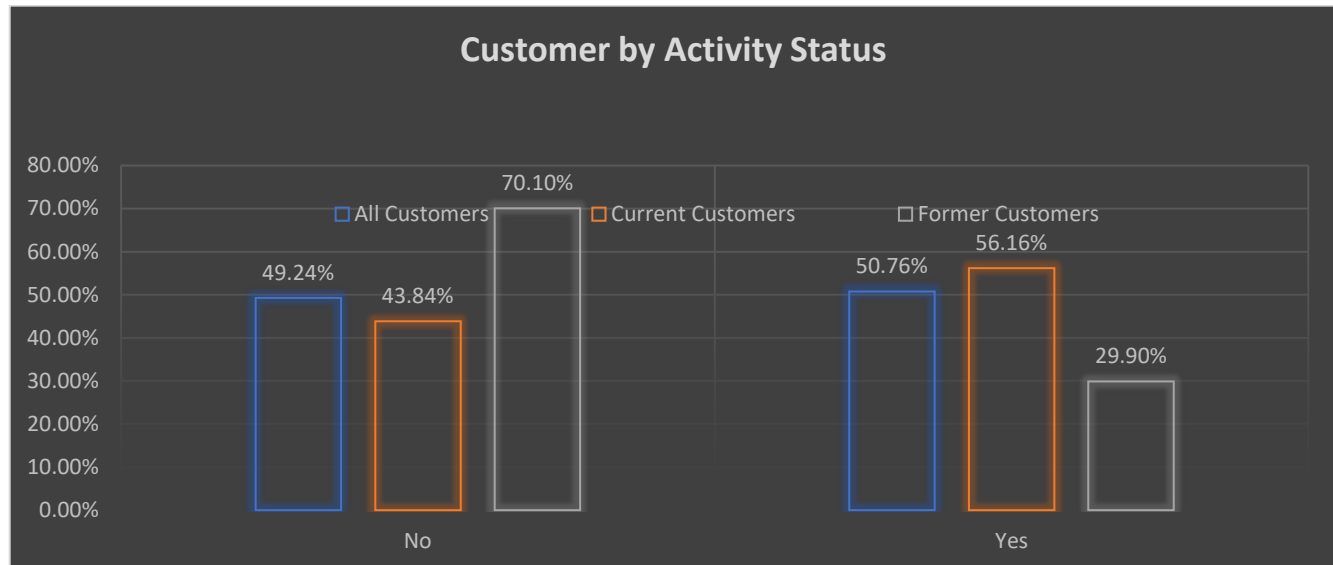
33.33% of customers ages 45-54 have left. 15.69% of customers ages 55-64 have left. Ages 25-44 are a smaller age group than those 40 and over, but Pig E. Bank has more customers in that age group.

Distribution of Age Groups for Former Customers

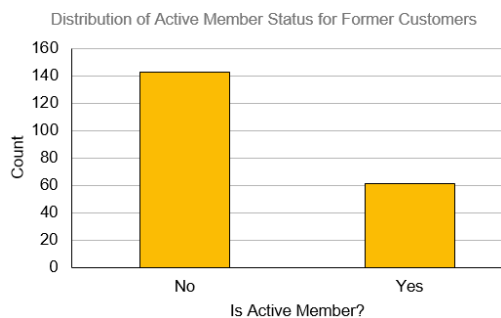
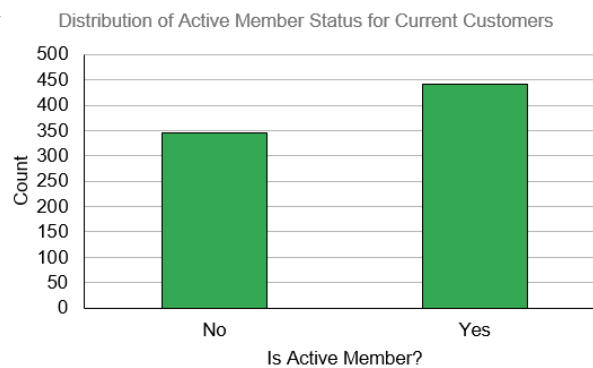
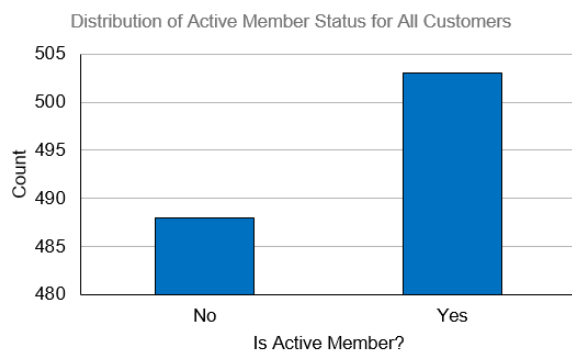


Exploring Variables: Customer Usage Activity Status

	All Customers	Current Customers	Former Customers
No	49.24%	43.84%	70.10%
Yes	50.76%	56.16%	29.90%



A significant proportion (70%) of former customers were inactive members, suggesting that activity status is a major factor in their decision to leave the bank.



I would consider activity status a factor for leaving the bank

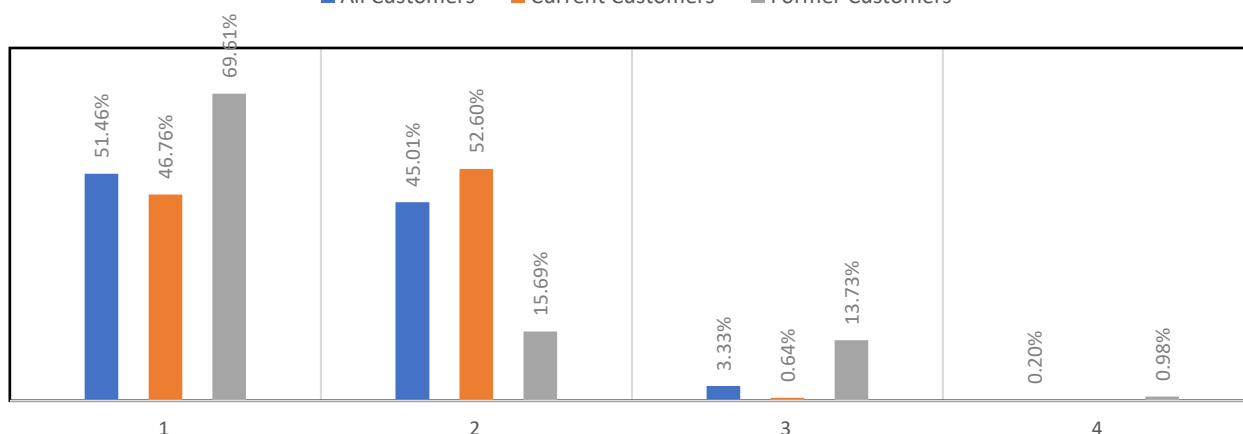
5.4: INTRO TO DATA MINING

Number of Products

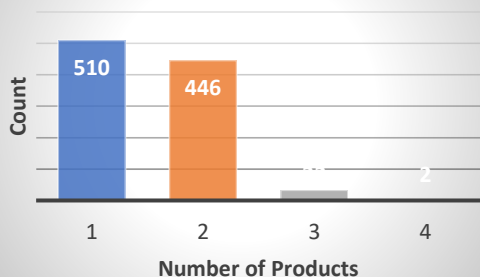
		All Customers	Current Customers	Former Customers
Number of Products	1	51.46%	46.76%	69.61%
	2	45.01%	52.60%	15.69%
	3	3.33%	0.64%	13.73%
	4	0.20%		0.98%

PERCENTAGE OF PRODUCTS

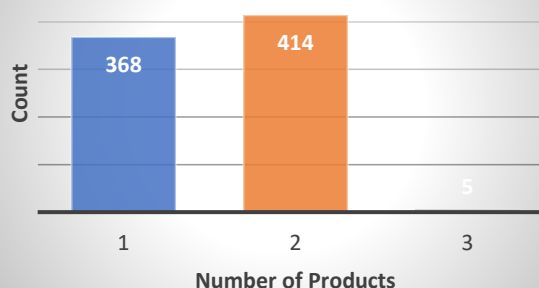
■ All Customers ■ Current Customers ■ Former Customers



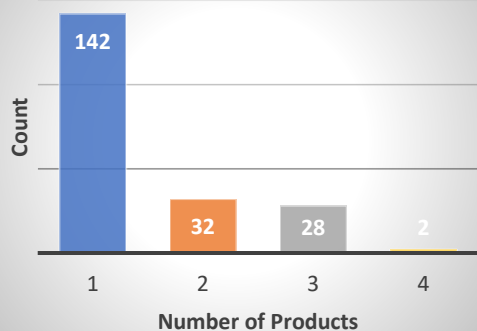
Distribution of Number of Products for All Customers



Distribution of Number of Products for Current Customers



Distribution of Number of Products for Former Customers



The majority of individuals who left the bank had only one product (70%), while among those who stayed, 47% had one product and 53% had two products. This disparity suggests that having multiple products may contribute to customer retention.

Credit Score

Credit score didn't seem to be a factor in leaving. The scores of those that left and those that stayed at the bank seemed to follow the same pattern

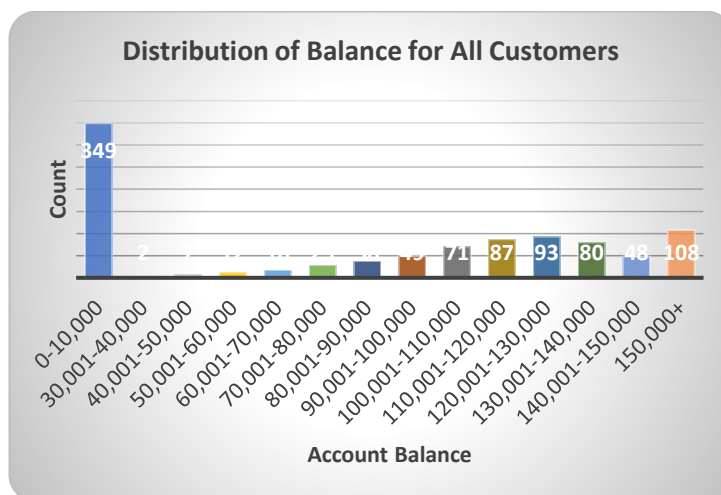
		All Customers	Current Customers	Former Customers	Notes
Credit Score	300-399	0.20%		0.98%	Credit scores are similar across groups.
	400-499	6.86%	6.61%	7.84%	
	500-599	24.12%	23.63%	25.98%	
	600-699	36.53%	36.59%	36.27%	
	700-799	26.44%	27.06%	24.02%	
	800-899	5.85%	6.10%	4.90%	

Tenure:

Tenure did not seem to be a factor in leaving. Those with both long and short terms left. The two groups also appear to be proportional to each other

		All Customers	Current Customers	Former Customers
Tenure	0	3.83%	4.07%	2.94%
	1	10.39%	9.28%	14.71%
	2	11.10%	10.55%	13.24%
	3	10.19%	10.29%	9.80%
	4	8.68%	8.77%	8.33%
	5	10.09%	10.17%	9.80%
	6	9.49%	9.28%	10.29%
	7	9.38%	9.91%	7.35%
	8	10.90%	11.18%	9.80%
	9	10.90%	11.05%	10.29%
	10	5.05%	5.46%	3.43%

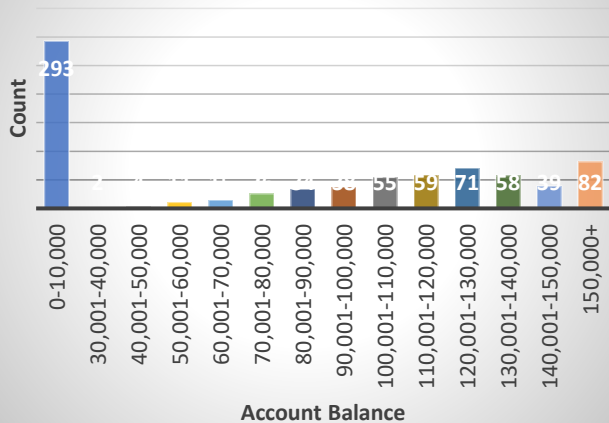
Account Balance:



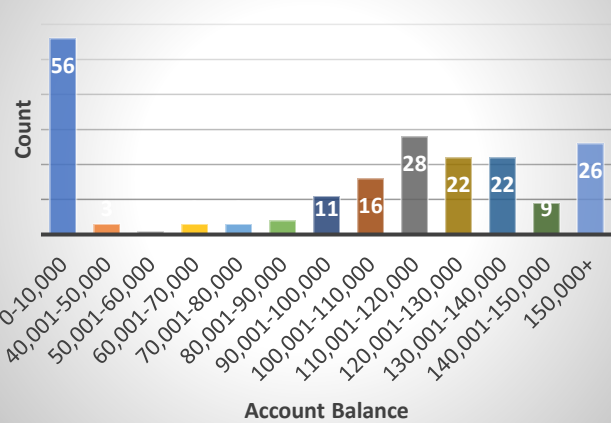
Insight: A larger proportion of the former customers have higher account balance (between 100k-140k and also balances more than 150k). This can also be connected to age (Older people having higher account balances)

Recommendation: Enhance personalized financial advisory services, focusing on wealth management, estate planning, and investment options tailored for their financial needs of people with higher balances.

Distribution of Balance for Current Customers



Distribution of Balance for Former Customers



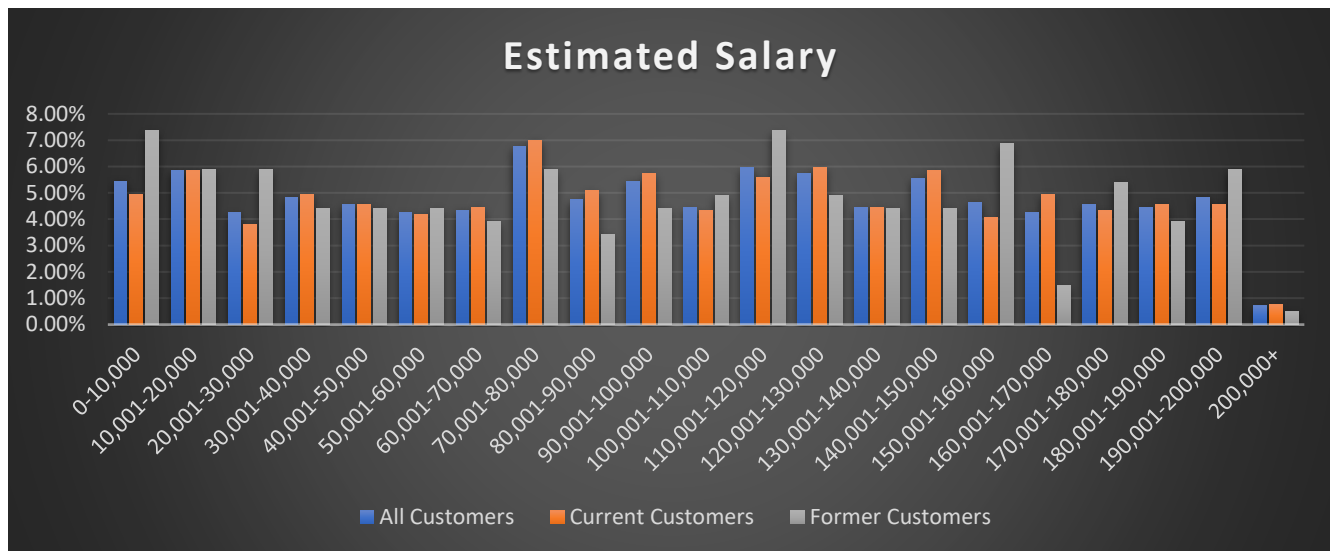
		All Customers	Current Customers	Former Customers
Account Balance	0	35.22%	37.23%	27.45%
	30,001-40,000	0.20%	0.25%	1.47%
	40,001-50,000	0.71%	0.51%	1.47%
	50,001-60,000	1.21%	1.40%	0.49%
	60,001-70,000	1.82%	1.91%	1.47%
	70,001-80,000	2.93%	3.30%	1.47%
	80,001-90,000	3.83%	4.32%	1.96%
	90,001-100,000	4.94%	4.83%	5.39%
	100,001-110,000	7.16%	6.99%	7.84%
	110,001-120,000	8.78%	7.50%	13.73%
	120,001-130,000	9.38%	9.02%	10.78%
	130,001-140,000	8.07%	7.37%	10.78%
	140,001-150,000	4.84%	4.96%	4.41%
	150,000+	10.90%	10.42%	12.75%

The percentage of the former customers is higher in the 110,001-120,000 (13.73%), 120,001-130,000 (10.78%), and 130,001-140,000 (10.78%) ranges.

Estimate Salary:

Estimated salaries are similar across groups

Estimated salary did not seem to be a factor in leaving the bank. For all salary ranges, only 0.49% - 7.35% of all customers left the bank. When looking at the percentage of customers leaving versus the total number of customers in a specific salary range, no group had an unusually high percentage of those that left. This is also true if we group the salary ranges together. 54.9% of customers with salaries of \$0-\$100,000 left, and 4591% of customers with salaries of \$100,000-\$199,999 left.



		All Customers	Current Customers	Former Customers
Estimated Salary	0-10,000	5.45%	4.96%	7.35%
	10,001-20,000	5.85%	5.84%	5.88%
	20,001-30,000	4.24%	3.81%	5.88%
	30,001-40,000	4.84%	4.96%	4.41%
	40,001-50,000	4.54%	4.57%	4.41%
	50,001-60,000	4.24%	4.19%	4.41%
	60,001-70,000	4.34%	4.45%	3.92%
	70,001-80,000	6.76%	6.99%	5.88%
	80,001-90,000	4.74%	5.08%	3.43%
	90,001-100,000	5.45%	5.72%	4.41%
	100,001-110,000	4.44%	4.32%	4.90%
	110,001-120,000	5.95%	5.59%	7.35%
	120,001-130,000	5.75%	5.97%	4.90%
	130,001-140,000	4.44%	4.45%	4.41%
	140,001-150,000	5.55%	5.84%	4.41%
	150,001-160,000	4.64%	4.07%	6.86%
	160,001-170,000	4.24%	4.96%	1.47%
	170,001-180,000	4.54%	4.32%	5.39%
	180,001-190,000	4.44%	4.57%	3.92%
	190,001-200,000	4.84%	4.57%	5.88%
	200,000+	0.71%	0.76%	0.49%

Conclusion and Recommendations:

When looking at customer demographics, there's always the possibility of bias. Pig E. Bank's products and services may appeal more to younger customers, or advertising of those products and services may not be reaching an older customer base. Older customers tend to prefer in-person interactions at branches, safer investments options, and may need retirement advisors or long-term care planning.

Recommendations include:

- We can survey those identified as most likely to leave the bank to get their input on why.
- To retain customers with higher balances we can look at expanding our investment and savings products. A person with a high balance may be saving for a large purchase or looking to invest. In both cases rate shopping would be a factor in where they keep their money.
- To increase activity, we should review the status of our long-term products. Mortgages, large installment loans, even direct deposits will keep customers at Pig E. Bank.
- To increase the number of products per customer, rewards and goals can be used to incentivize staff to promote products. We can also launch rate specials and targeted ads to increase customer engagement.

5.4: INTRO TO DATA MINING

Customer Churn Prediction: Factors Influencing Likelihood to Stay or Leave