Credit Risk Migration

Steven Esposito

Importing the datasets and merging them

IMPORT

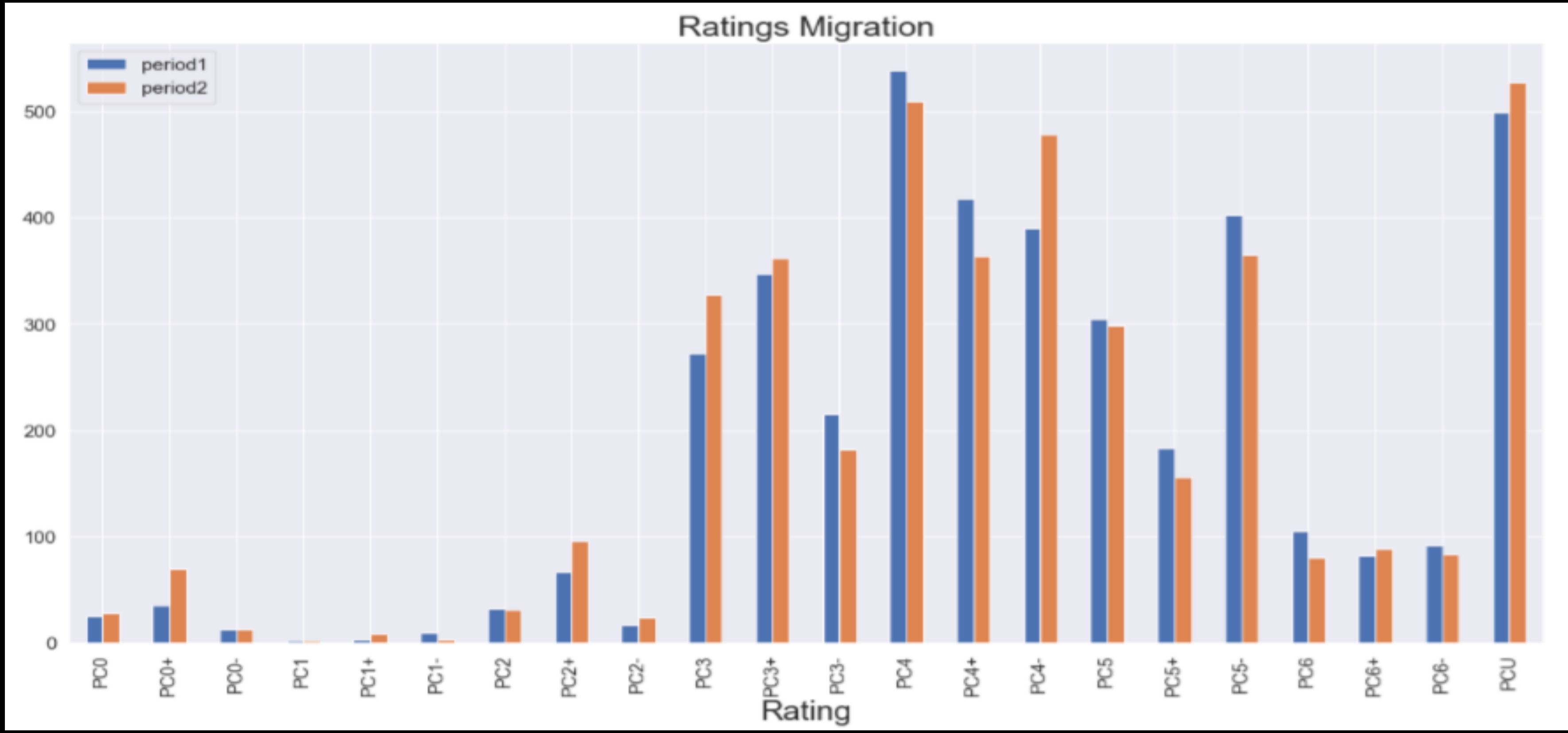
- period1 = pd.read_excel('Migration Case Data.xlsx', sheet_name='Period 1')
- period2 = pd.read_excel('Migration Case Data.xlsx', sheet_name='Period 2')

MERGE

- merged_inner = period1.merge(period2, on='ID', how='inner') → *common customers from both periods*
- merged_right = period1.merge(period2, on='ID', how='right') → *common customers plus new customers in period 2*
- merged_outer = period1.merge(period2, on='ID', how='outer') → *all customers from both periods*

Rating migration (all customers)

Excluding a couple of exceptions, it looks like there has been an increase in low rating and a decrease in high ratings. Since the Probability of Default is inversely proportional to the rating, this would suggest that from period 1 to period 2 there has been an increase in PD



Increased Risk Weight

NEW COLUMN FOR IRW

- merged_outer['IRW_1']= merged_outer['RWA_x'] / merged_outer['EAD Amount_x']
- merged_outer['IRW_2']= merged_outer['RWA_y'] / merged_outer['EAD Amount_y’]

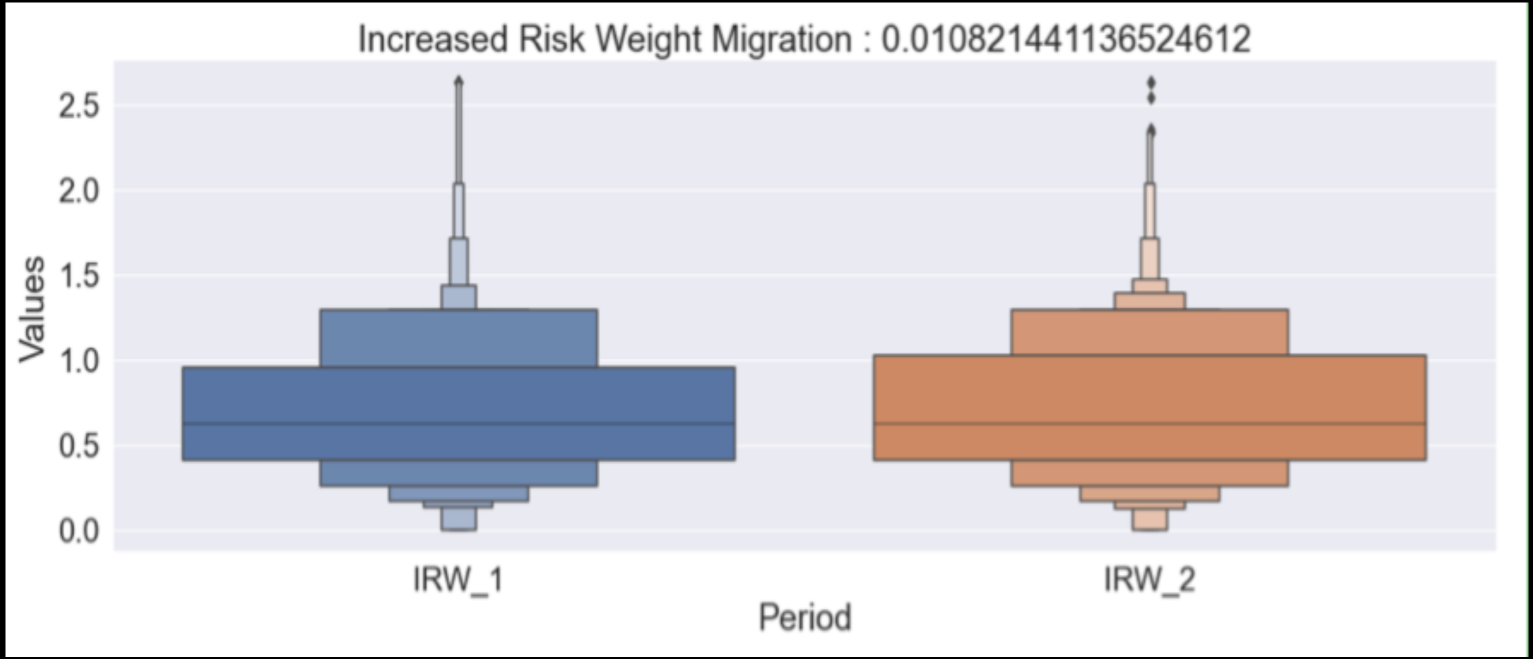
SEPARATE DATASET FOR IRW

- IRW_outer = merged_outer[['ID','IRW_1','IRW_2’]]
- IRW_outer = pd.melt(IRW_outer, id_vars="ID", var_name="Period", value_name="Values")

Increased Risk Weight

CALCULATING THE MEAN DIFFERENCE AND PLOTTING THE DISTRIBUTION

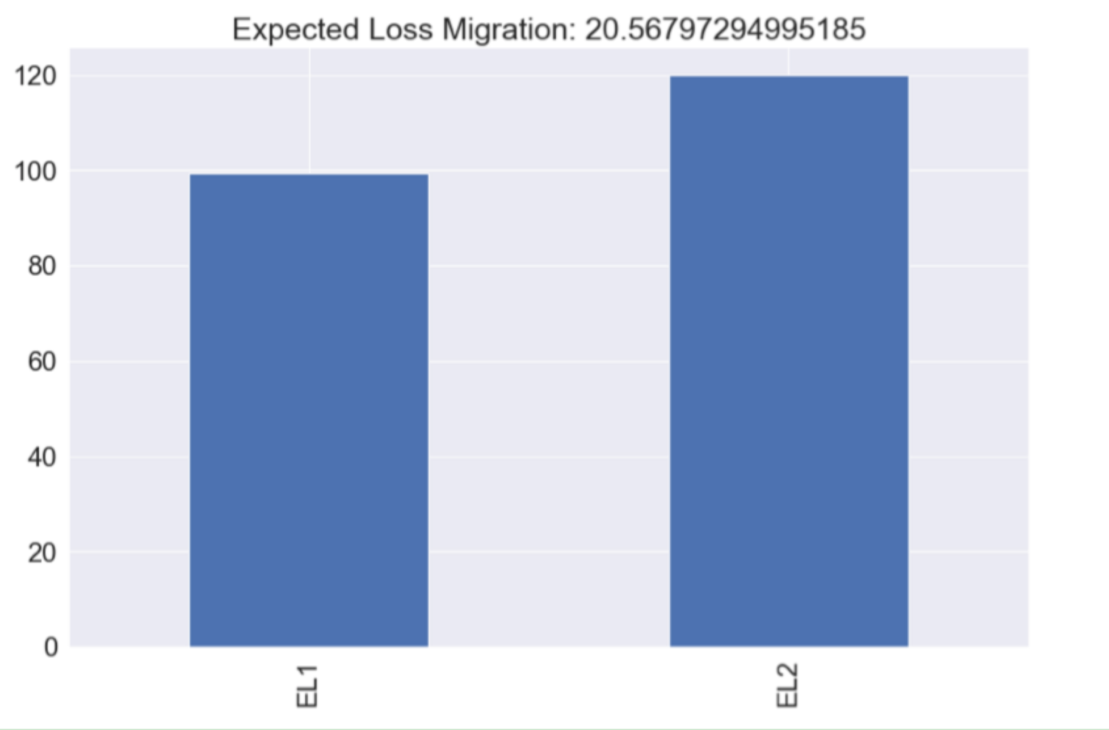
```
IRW_diff = merged_outer.IRW_2.mean()- merged_outer.IRW_1.mean()
g = sns.catplot(x="Period", y="Values",data=IRW_outer, kind="boxen", height=6, aspect=2.6);
plt.title(f'Increased Risk Weight Migration : {IRW_diff}', fontsize=25)
plt.show()
```



Expected Loss

CALCULATING THE MEAN DIFFERENCE AND PLOTTING THE RESULT

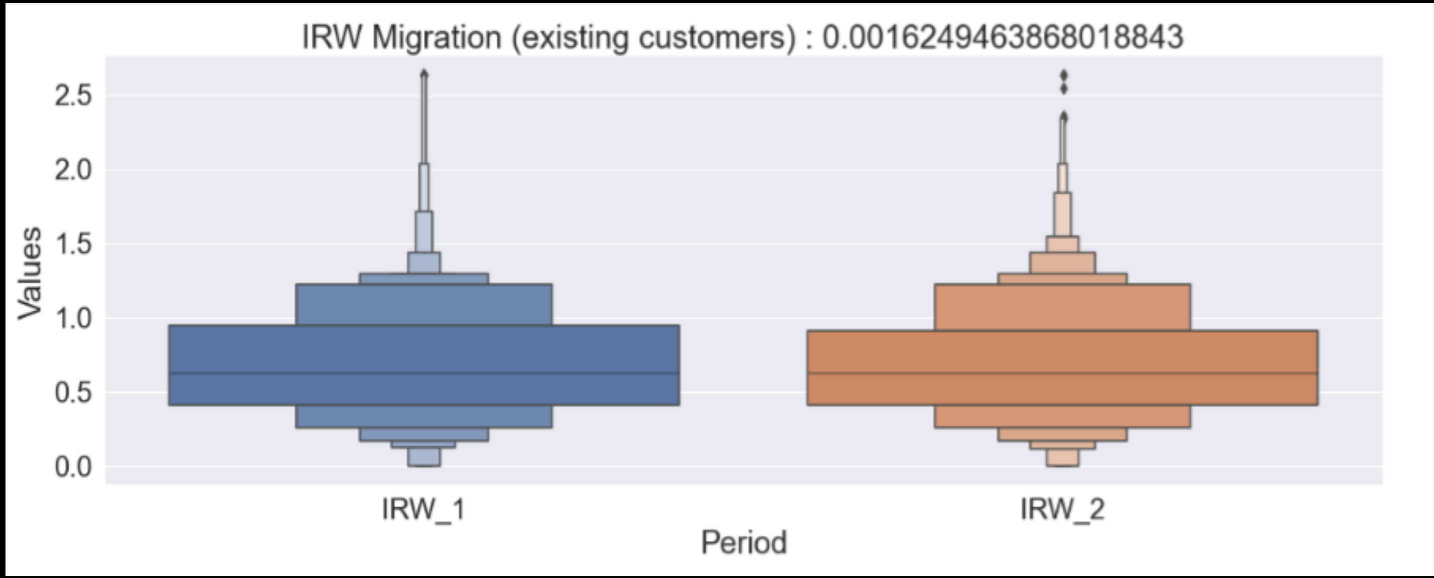
- EL_outer = merged_outer[['ID','Expected Loss_x','Expected Loss_y']]
- EL_outer = EL_outer.rename(columns={'Expected Loss_x':'EL1','Expected Loss_y':'EL2'})
- EL_diff = EL_outer.EL2.mean()- EL_outer.EL1.mean()
- EL_cols = ['EL1', 'EL2']
- EL_outer[EL_cols].mean().plot(kind='bar',figsize=(14,9))
- plt.title(f'Expected Loss Migration: {EL_diff}', fontsize=25)

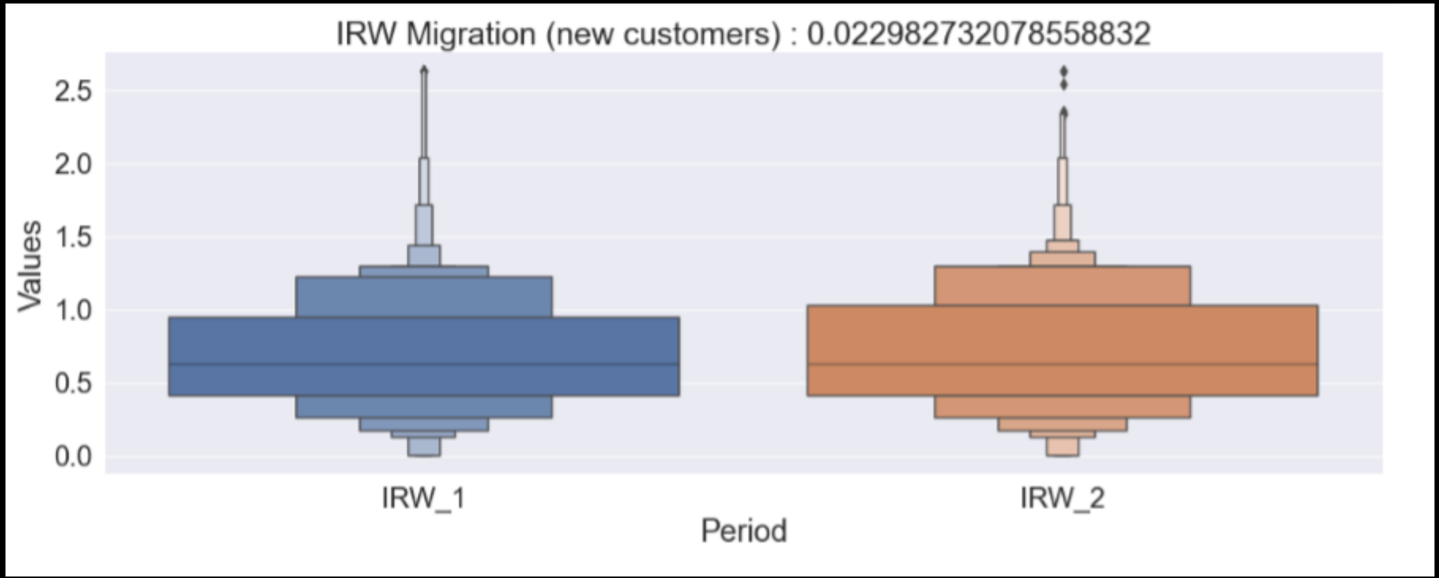


Impact of new customers VS existing customers

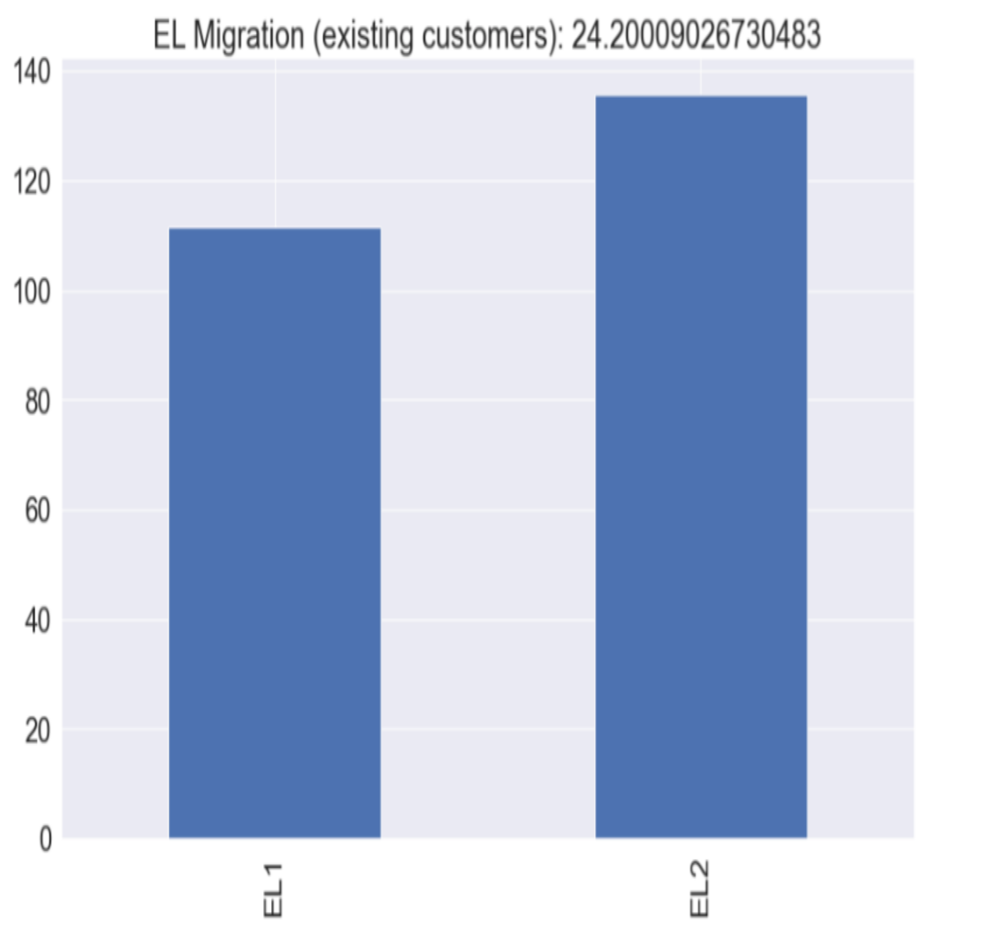
- So far, I have analyzed data taking into account all the customers: the ones that appear in both periods and the ones appearing only in one of them. How to measure the impact of new customers and existing customers?
- I will do this by comparing metrics between the 'merged_inner' set (containing common customers between the two periods) and the 'merged_right' one (containing common customers plus the new customers that appear in period 2).

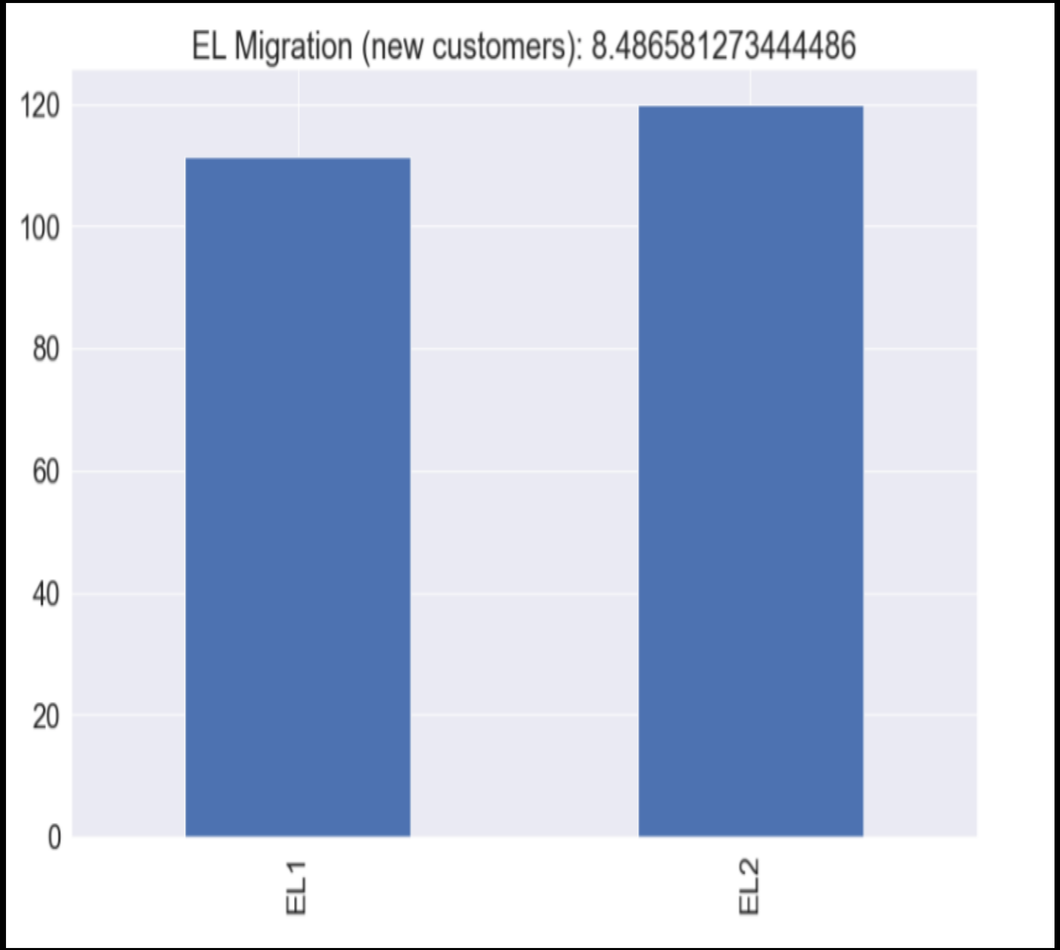
Increased Risk Weight



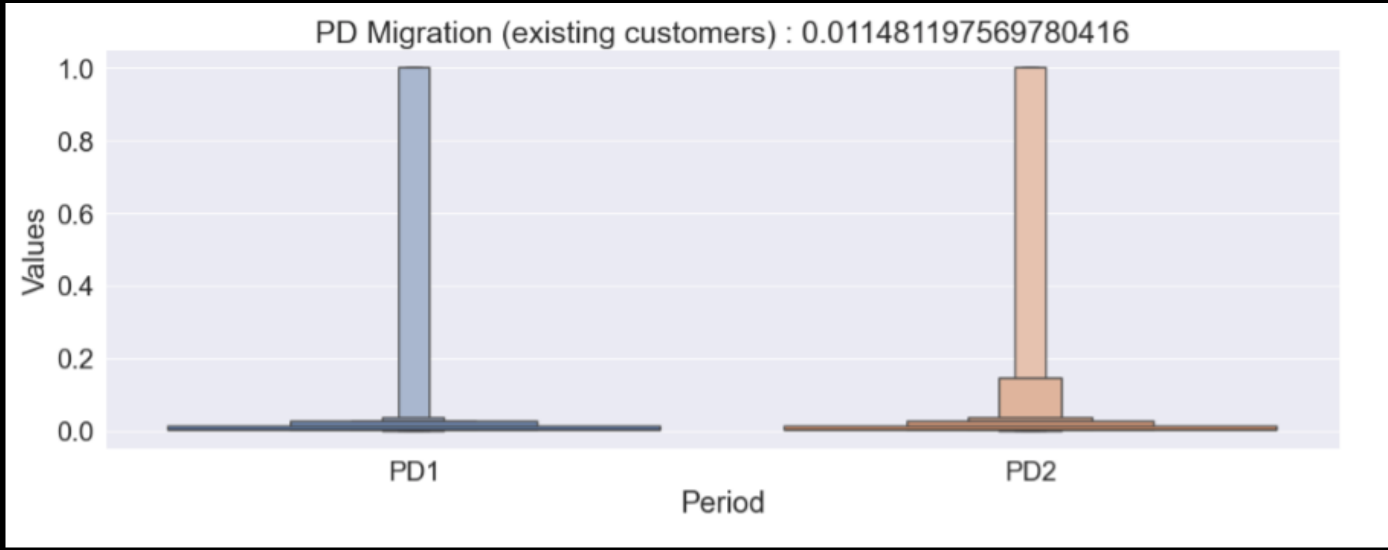


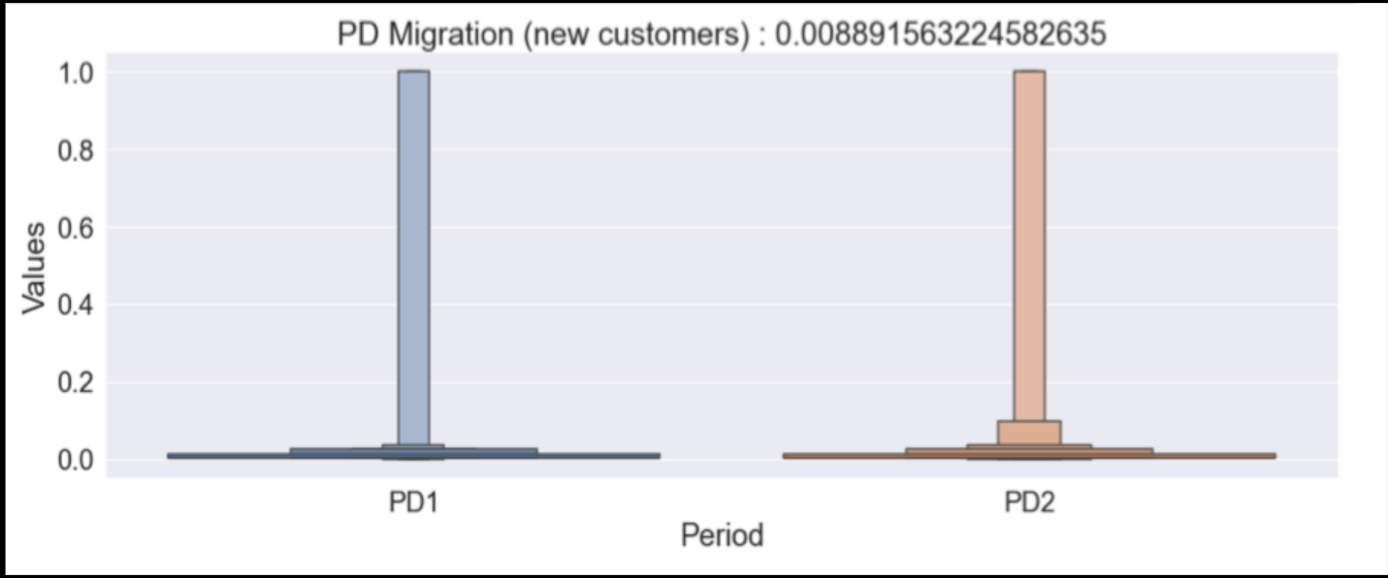
Expected Loss





Probability of Default

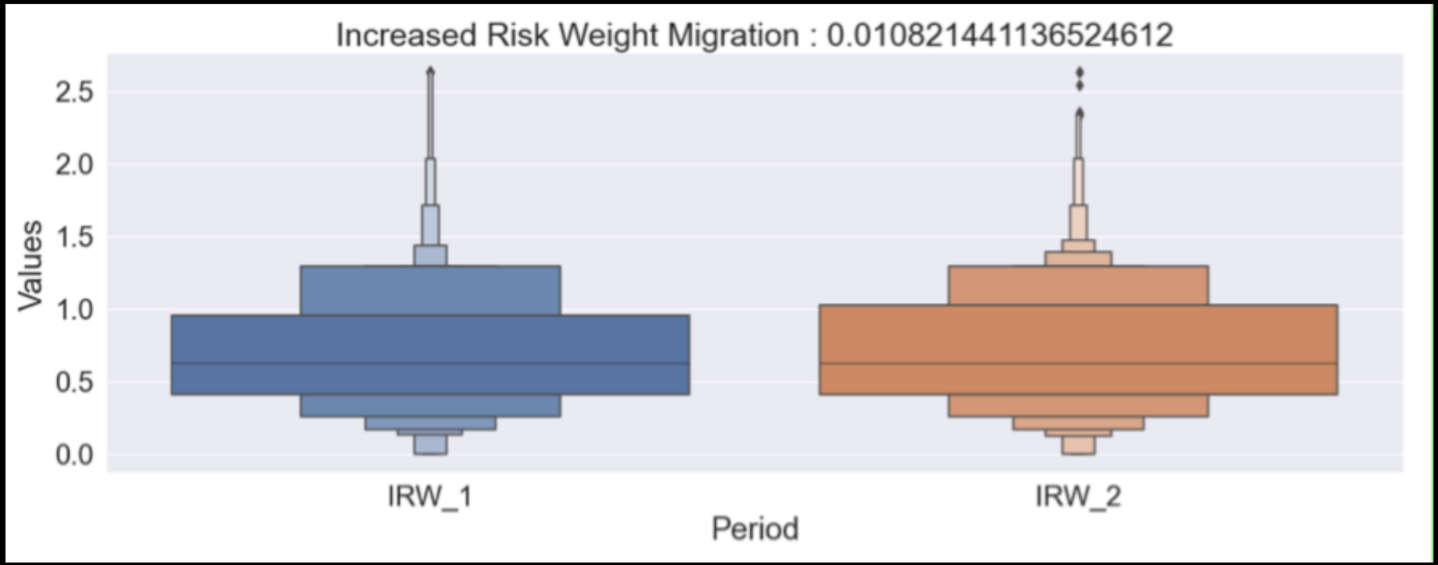


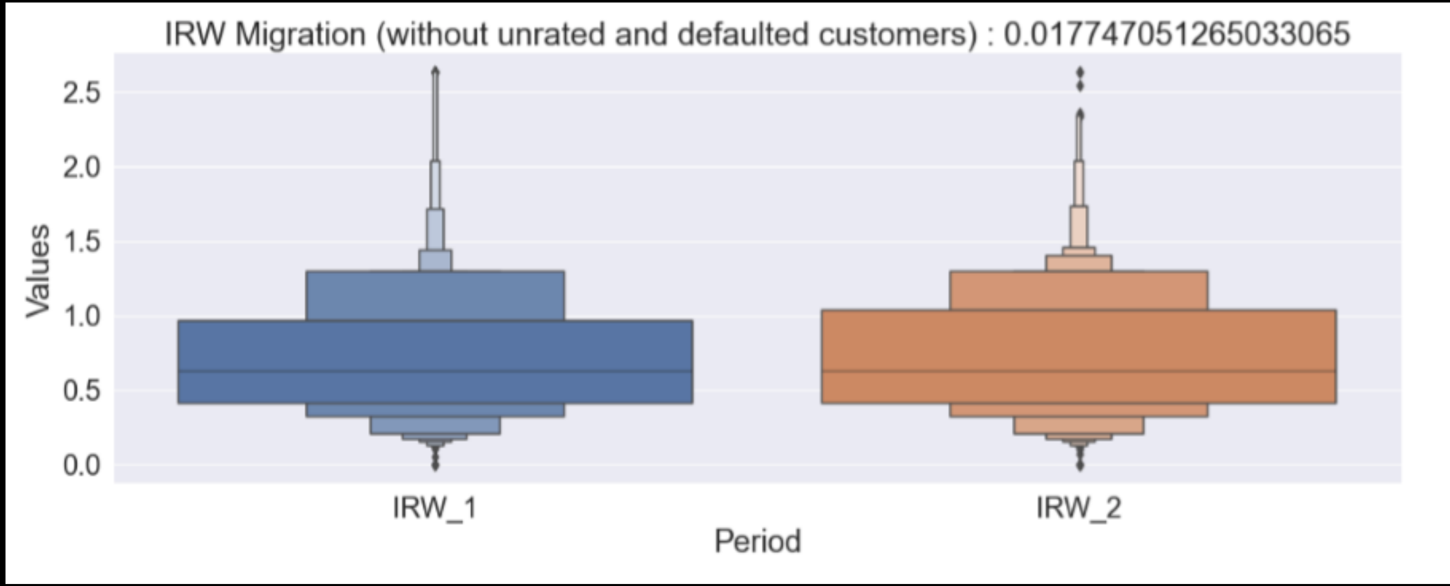


Impact of unrated and defaulted customers

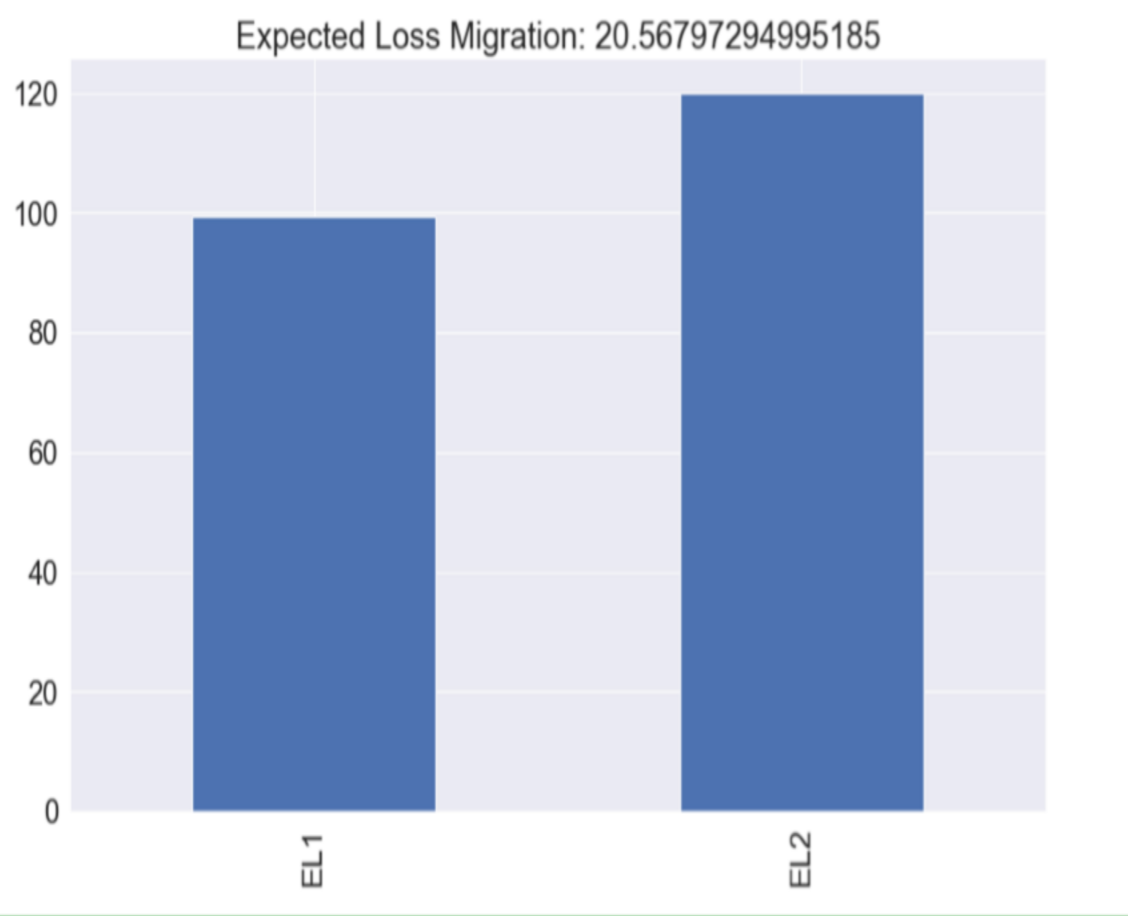
- The previous graphs showed that new customers had a bigger impact on the risk weight, however, it’s the existing customers that seem to influence the rise in PD and EL the most.
- What about the impact of unrated and defaulted customers? In order to answer this, I will first create a new dataset by dropping unrated and defaulted customers, then compare metrics with the dataset that contains all customers from both periods.

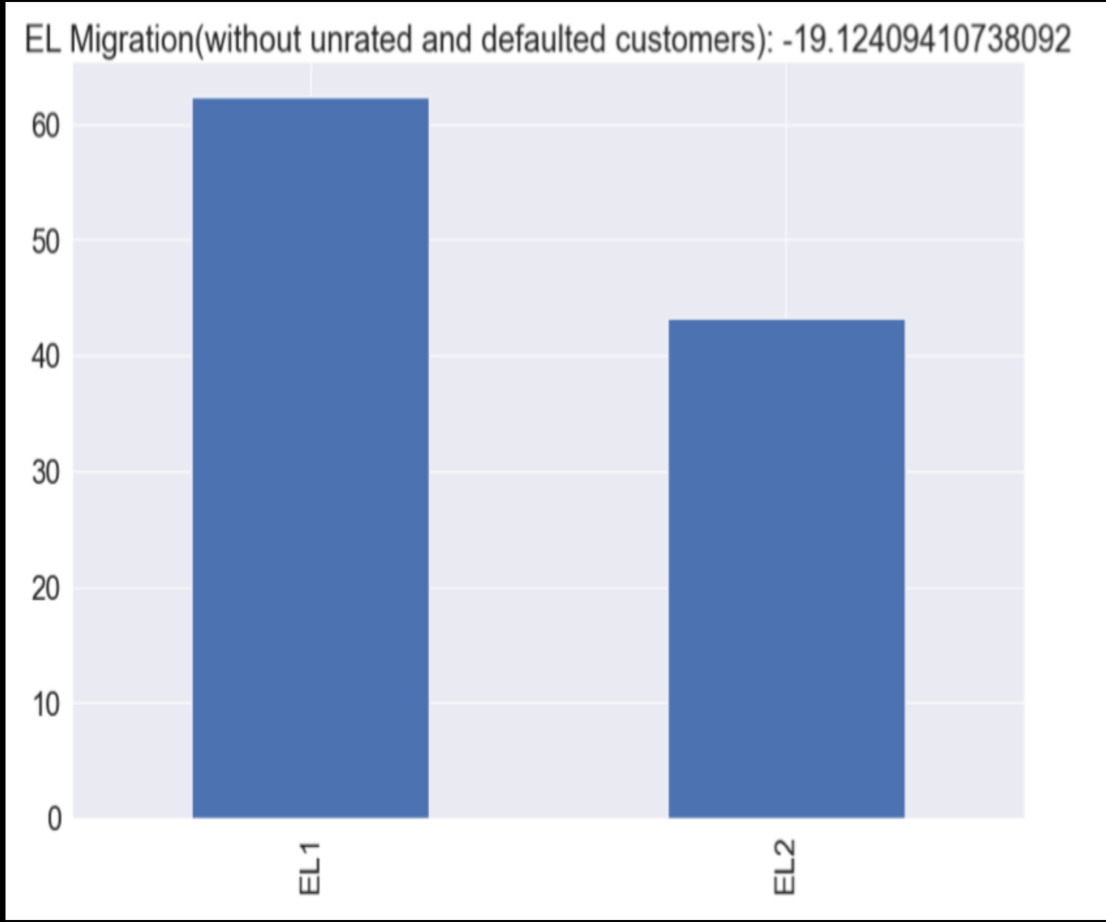
Increased Risk Weight



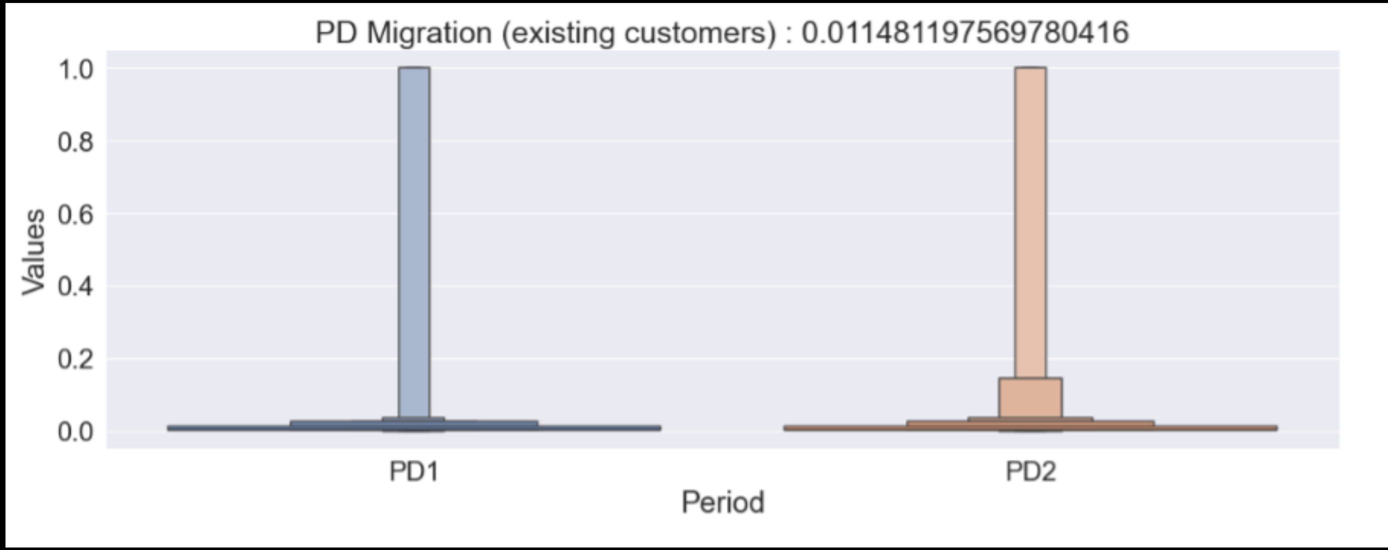


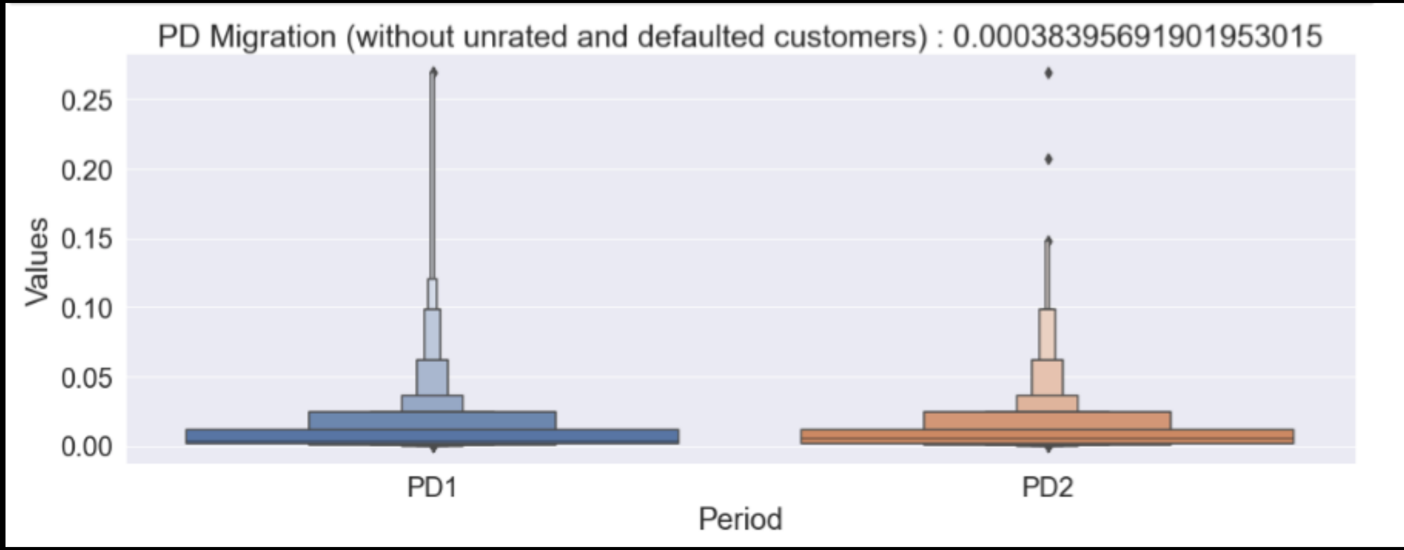
Expected Loss





Probability of Default





Impact of unrated and defaulted customers

- Unrated and defaulted customers seem to have a considerable impact on EL and PD. Without them, these two metrics would be considerably lower.

Which customers have increased/decreased capital consumption the most

- In order to answer this question, I will create a new column that will register the difference in capital requirements between period1 and period2. I will do my analysis considering the existing customers but excluding unrated and defaulted customers; for this I will first create a new dataframe 'merged_drop_inner'.

CREATION OF A NEW DATASET AND NEW COLUMN

- merged_drop_inner = period1_drop.merge(period2_drop, on='ID', how='inner')
- merged_drop_inner['Capital Requirement_diff'] = merged_drop_inner['Capital Requirement_y'] - merged_drop_inner['Capital Requirement_x']

CREATING A SEPARATE DATASET FOR CAPITAL REQ AND SORTING VALUES BY 'Capital Requirement_diff'

- capital_df = merged_drop_inner[['ID','Capital Requirement_x','Capital Requirement_y','Capital Requirement_diff']]
- capital_df = capital_df.sort_values(by=['Capital Requirement_diff'])

	ID	Capital Requirement_x	Capital Requirement_y	Capital Requirement_diff
2707	3396	839948.80305	91393.46090	-748555.34215
2708	3397	20919.01494	2400.85248	-18518.16246
1330	1706	27829.18251	9887.17751	-17942.00500
54	175	11853.98750	347.28761	-11506.69989
166	311	13794.12683	3300.40137	-10493.72546
...	...	...	...	...
22	135	37557.74841	45898.21387	8340.46546
958	1279	5884.93587	16342.82232	10457.88645
1298	1662	6.89405	12599.25967	12592.36562
17	130	13440.04953	32099.77193	18659.72240
532	757	10962.25338	38438.60465	27476.35127

Customers 3396, 3397,1706,175,311 have the lowest difference in Capital Requirement across the two periods, while customers 135,1279,1662,130,757 show the highest difference.

Thank you!

Steven Esposito