Total Dataset --> 9366

Delete Empty Sold --> 2

Present --> 9364

Delete ‘E’-Pf-Category --> 2

Present --> 9362

Delete negative period of holding --> 106

Present --> 9256

Unique desk\_id (train)--> 3993

After Deleting Empty sold --> 3991

After Removing ‘E’ Pf-category --> 3989

After Deleting Negative Period of Holding --> 3895

|  |  |
| --- | --- |
| Feature 1 | Portfolio\_id |
| Portfolio\_id | Unique data, No missing data |
| Desk\_id | Handle(10-1). Single desk-id can handle maximum of 10 Portfolio\_id.  [Only 2 desk\_id handles 10 Portfolio\_id]. |
| Office\_id | No of Office\_ID =2  OFF1 --> 2803 OFF2 --> 6453 |
| Pf\_Category | No of Pf\_Category = 4  A --> 2368 B --> 1209 C --> 5256 D --> 423 |
| Start\_date | Started\_Date Years with the portfolio count.  2004 - 66 2005 - 194 2006 - 210 2007 - 251 2008 - 631 2009 - 845  2010 - 937 2011 - 1569 2012 - 1181 2013 - 630 2014 - 692  2015 - 750 2016 - 795 2017 - 505 |
| Sold |  |
| Country\_code | Total No of Country = 5  M --> 5301 N --> 558 T --> 2906 U --> 61 Z --> 430 |
| Euribor\_rate |  |
| Currency | Total No of Currency = 5  CHF --> 5301 EUR --> 430 GBP --> 558 JPY --> 61 USD --> 2906 |
| Libor\_rate |  |
| Bought |  |
| Creation\_date | Creation\_Date Years with the portfolio count.  2004 - 66 2005 - 194 2006 - 210 2007 - 251 2008 - 631 2009 - 845  2010 - 937 2011 - 1570 2012 - 1181 2013 - 630 2014 - 691  2015 - 750 2016 - 796 2017 - 504 |
| Indicator\_code | Two Types of indicator code:  No\_in\_code --> 5589  TRUE --> 3667 |
| Sell\_date | Sell\_date Years with the portfolio count.  2004 - 49 2005 - 196 2006 - 202 2007 - 249 2008 - 612 2009 - 831  2010 - 919 2011 - 1626 2012 - 1156 2013 - 616 2014 - 678  2015 - 765 2016 - 763 2017 - 594 |
| Type | Total types = 8  A - 2269 B - 636 C - 558 D - 45 E - 5679 F - 61 G - 1 H - 7 |
| Hedge\_value | Three types of hedge\_value  No\_hedge\_value - 5591  FALSE - 2437  TRUE - 1228 |
| Status | Two Types of Status:  No\_status --> 2974  TRUE --> 6282 |
| Absolute Return |  |
| Actual Return(G/B) |  |
| Period\_of\_holding |  |
| Return |  |

|  |  |
| --- | --- |
| Feature 2 | Desk\_id |
| Portfolio\_id | Handle(10-1). Single desk-id can handle maximum of 10 Portfolio\_id.  [Only 2 desk\_id handles 10 Portfolio\_id]. |
| Desk\_id | Presence of repeated data. Presence of missing data. Presence of missing data, filled with no\_desk\_id. [No.of.missing data : 3665].  No.of unique data including no\_desk\_id : 3895 |
| Office\_id | No of unique desk\_id in each office\_id  OFF1 --> 2140 [no\_desk\_id = 381]  OFF2 --> 2082 [no\_desk\_id = 3284] |
| Pf\_Category | Types of pf\_category with count of unique desk\_id:  A - 1864 [no\_Desk\_id - 381] B - 418 [no\_Desk\_id - 754]  C - 1756 [no\_Desk\_id - 2530] D - 265 [no\_Desk\_id - 0] |
| Start\_date | Started\_Date Years with the Desk\_id unique counts.  2004 - 61 2005 - 180 2006 - 210 2007 - 248 2008 - 567 2009 - 761  2010 - 635 2011 - 778 2012 - 455 [no\_desk\_id - 293]  2013 - 1 [no\_desk\_id - 630] 2014 - 1 [no\_desk\_id - 692]  2015 - 1 [no\_desk\_id - 750] 2016 - 1 [no\_desk\_id - 795]  2017 - 1 [no\_desk\_id - 505] |
| Sold |  |
| Country\_code | No of country\_code with count of unique desk\_id:  M - 1775 [no\_Desk\_id - 2539] N - 340 [no\_desk\_id - 207]  T - 1883 [no\_desk\_id - 858] U - 8 [no\_desk\_id - 54]  Z - 266 [no\_desk\_id - 7] |
| Euribor\_rate |  |
| Currency | Currency with count of unique desk\_id:  CHF - 1775 [no\_Desk\_id - 2539] GBP - 340 [no\_desk\_id - 207]  USD - 1883 [no\_desk\_id - 858] JPY - 8 [no\_desk\_id - 54]  ZUR - 266 [no\_desk\_id - 7] |
| Libor\_rate |  |
| Bought |  |
| Creation\_date | Creation\_Date Years with the Desk\_id unique counts.  2004 - 61 2005 - 180 2006 - 210 2007 - 248 2008 - 567 2009 - 761  2010 - 635 2011 - 779 2012 - 454 [no\_desk\_id - 294]  2013 - 1 [no\_desk\_id - 630] 2014 - 1 [no\_desk\_id - 691]  2015 - 1 [no\_desk\_id - 750] 2016 - 1 [no\_desk\_id - 796]  2017 - 1 [no\_desk\_id - 504] |
| Indicator\_code | Indicator\_code with unique desk id  No\_in\_code - 3892 [no\_desk\_id - 0]  TRUE - 3 [no\_desk\_id - 3665] |
| Sell\_date | Sell\_date Years with the Desk\_id unique counts.  2004 - 44 2005 - 183 2006 - 201 2007 - 247 2008 - 548 2009 - 761 2010 - 626 2011 - 815  2012 - 470 [no\_desk\_id - 249] 2013 - 1 [no\_desk\_id - 616] 2014 - 1 [no\_desk\_id - 678] 2015 - 1 [no\_desk\_id - 765]  2016 - 1 [no\_desk\_id - 763] 2017 - 1 [no\_desk\_id - 594] |
| Type | Type and unique desk\_id  A - 1773 [no\_desk\_id - 380] B - 153 [no\_desk\_id - 477]  C - 340 [no\_desk\_id - 207] D - 37 [no\_desk\_id - 9]  E - 2021 [no\_desk\_id - 2530] F - 8 [no\_desk\_id - 54]  G - 1 [no\_desk\_id - 1] H - 1 [no\_desk\_id -7] |
| Hedge\_value | Hedge\_value and unique desk\_id  No\_hedge\_value - 3894 [no\_desk\_id - 1]  FALSE - 1 [no\_desk\_id - 2437]  TRUE - 2 [no\_desk\_id -1227] |
| Status | Status with unique desk\_id  No\_status - 2565 [no\_desk\_id - 23]  TRUE - 1331 [no\_desk\_id - 3642] |
| Absolute Return |  |
| Actual Return(G/B) |  |
| Period\_of\_holding |  |
| Return |  |

|  |  |
| --- | --- |
| Feature 3 | Office\_id |
| Portfolio\_id | The no of portfolio\_id in each office\_id:  No of Office\_ID = 2  OFF1 --> 2803 OFF2 --> 6453 |
| Desk\_id | No of unique desk\_id in each office\_id  OFF1 --> 2140 [no\_desk\_id = 381]  OFF2 --> 2082 [no\_desk\_id = 3284] |
| Office\_id | Presence of repeated values, No missing values.  No of Office\_ID =2  OFF1 --> 2803 OFF2 --> 6453 |
| Pf\_Category | No of Pf\_Category = 4  OFF1 --> A , B , C , D  A - 2368 , B - 1 , C - 11 , D - 423  OFF2 --> B , C  B - 1208 , C - 5245 |
| Start\_date | --> 2007 contains only OFF1  --> 2004 - 2008 [OFF1 > OFF2] - OFF1 is not so high  --> 2008 - 2017 [OFF1 < OFF2] - OFF2 is very high  --> 2011 contains highest OFF1 and OFF2 |
| Sold |  |
| Country\_code | Total No of Country = 5  OFF1 --> M , N , T , Z  M - 14 , N - 95 , T - 2271 , Z - 423  OFF2 --> M , N , T , U , Z  M - 5287 , N - 463 , T - 635 , U - 61 , Z - 7 |
| Euribor\_rate |  |
| Currency | Total No of Currency = 5  OFF1 --> CHF , GBP , USD , EUR  CHF - 14 , GBP - 95 , USD - 2271 , EUR - 423  OFF2 --> CHF , GBP , USD , JPY , EUR  CHF - 5287 , GBP - 463 , USD - 635 , JPY - 61 , EUR - 7 |
| Libor\_rate |  |
| Bought |  |
| Creation\_date |  |
| Indicator\_code | Two Types of indicator code:  OFF1 --> No\_in\_code , TRUE  No\_in\_code - 2422 , TRUE --> 381  OFF2 --> No\_in\_code , TRUE  No\_in\_code - 3167 , TRUE --> 3286 |
| Sell\_date |  |
| Type | Total types = 8  OFF1 --> A , B , C , D , E , G  A - 2269 B - 1 C - 95 D - 3 E - 434 G - 1  OFF2 --> B , C , D , E , F , H  B - 635 , C - 463 , D - 42 , E - 5245 , F - 61 , H - 7 |
| Hedge\_value | Three types of hedge\_value  OFF 1 --> No\_hedge\_value, TRUE , FALSE  No\_hedge\_value - 2423 , FALSE - 287 , TRUE - 93  OFF 2 --> No\_hedge\_value, TRUE , FALSE  No\_hedge\_value - 3168 , FALSE - 2150 , TRUE - 1135 |
| Status | Two Types of Status:  OFF1 --> No\_status , TRUE  No\_status --> 1886 , TRUE --> 917  OFF2 --> No\_status , TRUE  No\_status --> 1088 , TRUE --> 5365 |
| Absolute Return |  |
| Actual Return(G/B) |  |
| Period\_of\_holding |  |
| Return |  |

|  |  |
| --- | --- |
| Feature 4 | Pf\_Category |
| Portfolio\_id | The no of portfolio\_id in each Pf\_category:  No of Pf\_category = 4  A --> 2368 B --> 1209 C --> 5256 D --> 423 |
| Desk\_id | No of unique desk\_id in each Pf\_Category  A --> 1864 [no\_desk\_id = 381]  B --> 418 [no\_desk\_id = 754]  C --> 1756 [no\_desk\_id = 2530]  D --> 265 [no\_desk\_id = 0] |
| Office\_id | Number of office\_id in each pf\_category  A --> OFF1 - 2368  B --> OFF1 - 1 , OFF2 - 1208  C --> OFF1 - 11 , OFF2 - 5245  D --> OFF1 - 423 |
| Pf\_Category | The values are repeated and no missing data.  No of Pf\_Category = 4  A --> 2368 B --> 1209 C --> 5256 D --> 423 |
| Start\_date | A --> 2004 - 2008 A is greater [2007 contains only A]  B --> Present in all year except[2007]. In the place of 2nd and 3rd  C --> 2009 -2017 C is greater [ 2011,2012 very high]  D --> It comes only in 2010,2011 [2010<2011] |
| Sold |  |
| Country\_code | Total No of Country = 5  A --> M - 3 , N - 95 , T - 2270  B --> M - 42 , N - 463 , T - 636 , U - 61 , Z - 7  C --> M -5256  D --> Z - 423 |
| Euribor\_rate |  |
| Currency | Total No of Currency = 5  A --> CHF - 3 , GBP - 95 , USD - 2270  B --> CHF - 42 , GBP - 463 , USD - 636 , JPY - 61 , EUR - 7  C --> CHF -5256  D --> EUR - 423 |
| Libor\_rate |  |
| Bought |  |
| Creation\_date |  |
| Indicator\_code | Two Types of indicator code:  A - no\_in\_code - 1987 , TRUE - 381  B - no\_in\_code - 454 , TRUE - 755  C - no\_in\_code - 2725 , TRUE - 2531  D - no\_in\_code - 423 |
| Sell\_date |  |
| Type | Total types = 8  A --> A - 2269 , C - 95 , D - 3 , G - 1  B --> B - 636 , C - 463 , D - 42 , F - 61 , H - 7  C --> E - 5256  D --> E - 423 |
| Hedge\_value | Three types of hedge\_value  A --> No\_hedge\_value - 1988 , TRUE - 93 , FALSE - 287  B --> No\_hedge\_value - 454 , TRUE - 190 , FALSE - 565  C --> No\_hedge\_value - 2726 , TRUE - 945 , FALSE - 1585  D --> No\_hedge\_value - 423 |
| Status | Two Types of Status:  A --> no\_status - 1452 , TRUE - 916  B --> no\_status - 276 , TRUE - 933  C --> no\_status - 823 , TRUE - 4433  A --> no\_status - 423 |
| Absolute Return |  |
| Actual Return(G/B) |  |
| Period\_of\_holding |  |
| Return |  |

|  |  |
| --- | --- |
| Feature 5 | Start\_date |
| Portfolio\_id | Started\_Date Years with the portfolio count.  2004 - 66 2005 - 194 2006 - 210 2007 - 251 2008 - 631 2009 - 845  2010 - 937 2011 - 1569 2012 - 1181 2013 - 630 2014 - 692  2015 - 750 2016 - 795 2017 - 505 |
| Desk\_id | Started\_Date Years with the Desk\_id unique counts.  2004 - 61 2005 - 180 2006 - 210 2007 - 248 2008 - 567 2009 - 761  2010 - 635 2011 - 778 2012 - 455 [no\_desk\_id - 293]  2013 - 1 [no\_desk\_id - 630] 2014 - 1 [no\_desk\_id - 692]  2015 - 1 [no\_desk\_id - 750] 2016 - 1 [no\_desk\_id - 795]  2017 - 1 [no\_desk\_id - 505] |
| Office\_id | Started\_Date Years with the Office\_id counts.  --> 2007 contains only OFF1  --> 2004 - 2008 [OFF1 > OFF2] - OFF1 is not so high  --> 2008 - 2017 [OFF1 < OFF2] - OFF2 is very high  --> 2011 contains highest OFF1 and OFF2 |
| Pf\_Category | Started\_Date Years with the Pf\_Category counts.  A --> 2004 - 2008 A is greater [2007 contains only A]  B --> Present in all year except[2007]. In the place of 2nd and 3rd  C --> 2009 -2017 C is greater [ 2011,2012 very high]  D --> It comes only in 2010,2011 [2010<2011] |
| Start\_date | The range of the year starts from 2004-2017. From May 2004 (not June) To October 2017. |
| Sold |  |
| Country\_code | Started\_Date Years with the Country code counts.  M --> Present in 2005,2006,[2008-2017]. High from [2009-2017]  N --> Present in all the year. Some what low in all the year.  T --> Present in all the year. Some what high from [2004-2008]. Avg[2009-2017]  U --> Present from [2011-2017]. Low in all the year.  Z --> Present in 2010,2011,2015 [high - 2011 low - 2010,2015]. |
| Euribor\_rate |  |
| Currency | Started\_Date Years with the Currency counts.  CHF --> Present in 2005,2006,[2008-2017]. High from [2009-2017]  GBP --> Present in all the year. Some what low in all the year.  USD --> Present in all the year. Some what high from [2004-2008]. Avg[2009-2017]  JPY --> Present from [2011-2017]. Low in all the year.  EUR --> Present in 2010,2011,2015 [high - 2011 low - 2010,2015]. |
| Libor\_rate |  |
| Bought |  |
| Creation\_date | In 1387 places creation\_date have been set before start\_date.  Negative days - 2007, 2009- 2017 [max:2010-2017] |
| Indicator\_code | 2004 - 2010 --> only no\_in\_code [counts get increases every year]  2011 , 2012 --> Both no\_in\_code and TRUE are present.  2011 - no\_in\_code > TRUE [no\_in\_code-very high,TRUE-very low]  2012 - no\_in\_code > TRUE [no\_in\_code & TRUE are moderate]  2013 - 2017 --> Only True [moderate range] |
| Sell\_date | Their is 100 negative values.  Negative values - 2012 - 2014 , 2016 |
| Type |  |
| Hedge\_value |  |
| Status |  |
| Absolute Return |  |
| Actual Return(G/B) |  |
| Period\_of\_holding |  |
| Return |  |