In [9]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

In [10]:
```
%matplotlib inline
```

In [30]:
```
df=pd.read_csv("C:\\Users\\Shay\\Documents\Yari\\Data analysis Python\\Data\\CCB_CI_UFO_HD
```

In [31]:
```
df = df[["HDDSNENDDATE","HDDSN","PFCODE","MFGID","HDDTRIAL","TESTCODE","TESTCODEC","QUALIF
        "LHD","PHD","AePort","FinalHeadTotalLBA","ACCHeadTotalLBA","AccHeadTotalLbaNoAti"
```

In [32]:
```
df=df.drop(index=df[df['QUALIFIER'] == 1020].index)
```

In [41]:
```
#Set the Target LBA per surface
#Each product has a different LBA Target
#First 3 letters of TestCode is product identifier
```

In [ ]:
```
# The if..else statement evaluates test expression and will execute the body of if only wh
#the test condition is True . If the condition is False , the body of else is executed.
#Indentation is used to separate the blocks.
```

In [33]:
```
if  df['TESTCODEC'].str[:3].any() == 'PCM':
    CmrTgtLba = 246658123
    SmrTgtLba = 279332228
elif (df['TESTCODEC'].str[:3].any()) == 'LXJ' or (df['TESTCODEC'].str[:3].any()) == 'XL9'
    CmrTgtLba = 246669158
    SmrTgtLba = 282846286
elif (df['TESTCODEC'].str[:3].any()) == 'PDQ' or (df['TESTCODEC'].str[:3].any()) == 'ADQ':
    CmrTgtLba = 273906721
    SmrTgtLba = 333662709
elif df['TESTCODEC'].str[:3].any() == 'TCL':
    CmrTgtLba = 246669158
    SmrTgtLba = 282846286
else:
    CmrTgtLba = 9999
    SmrTgtLba = 9999
```

In [43]:
```
#Create 2 new Column "iACC" and "pACC"
#Convert LBA's to ACC based on following equations
#CMR:
#iACC =(100 * AccHeadTotalLbaNoAti / 246669158)
#pACC =(100 * ACCHeadTotalLBA/ 246669158)
```

In [34]:
```
if     [df['QUALIFIER'].any == "100" or df['QUALIFIER'].any == 100 or df['QUALIFIER'].any
        df['iACC'] = df['AccHeadTotalLbaNoAti'] / CmrTgtLba * 100
        df['pACC'] = df['AccHeadTotalLbaFormat'] / CmrTgtLba * 100
else:
        [df['QUALIFIER'].any == "200" or df['QUALIFIER'].any == 200 or df['QUALIFIER'].any
        df['iACC'] = df['AccHeadTotalLbaFormat'] / SmrTgtLba * 100
        df['pACC'] = df['AccHeadTotalLbaFormat'] / SmrTgtLba * 100
```

In [36]:
```
if   [df['QUALIFIER'].any == "100" or df['QUALIFIER'].any == 100 or df['QUALIFIER'].any ==
     df['FinACC'] = df['FinalHeadTotalLBA'] / CmrTgtLba * 100
elif [df['QUALIFIER'].any == "200" or df['QUALIFIER'].any == 200 or df['QUALIFIER'].any ==
     df['FinACC'] = df['FinalHeadTotalLBA'] / SmrTgtLba * 100
```

```
olumn 'ACC_margin'
```

In [37]:
```
df['ACC_margin'] = (df['pACC'] - df['FinACC'])
```

In [38]:
```
df.drop(['FinalHeadTotalLBA', 'ACCHeadTotalLBA', 'AccHeadTotalLbaNoAti', 'AccHeadTotalLbaP
```

In [39]:
```
df.drop(['PHD'], axis=1, inplace=True)
```

In [40]:
```
df.rename(columns={'AEPORT':'PHD'}, inplace=True)
```

In [41]:
```
df.drop(['HDDSNENDDATE', 'TESTCODE','QUALIFIER','SubQualifier'], axis = 1, inplace = True)
```

In [42]:
```
df.to_csv('C:\\Users\\Shay\\Documents\Yari\\Data analysis Python\\Data\\ACC_PREPROCESS_Tes
```